Procedural Decision-Making In The Face Of Complexity

Gonzalo Arrieta* Kirby Nielsen†

Preliminary Draft - Please do not circulate

Click here to see latest version

October 4, 2023

Abstract

Decades of evidence on decision-making suggest that complexity is an essential factor influencing choices. Evidence is consistent with individuals using different choice *processes* in the face of complexity, but choice processes are inherently difficult to observe. We introduce an experimental methodology to test the hypothesis that individuals resort to more "procedural" decision processes as decisions get more complex. We characterize procedural decision-making as choice processes that are more describable, and we measure this by incentivizing other participants to guess what decision-makers chose based on decision-makers' own descriptions of their choice process. In two different domains and with two different notions of complexity, we show that procedural decision-making increases as we exogenously vary the complexity of the environment. We show evidence that procedural decision-making is a choice simplification that reduces the cognitive costs of decision-making and that using procedural choice processes can lead decision-makers to make different choices and of different quality.

I. Introduction

A large body of work in economics, psychology, and beyond studies the choices that individuals make and the preferences that these choices reveal. Work in many subfields has uncovered rich data on the shape and structure of individuals' preferences.[1] However, relatively

*garrieta@stanford.edu; Department of Economics, Stanford University

†kirby@caltech.edu; Division of the Humanities and Social Sciences, California Institute of Technology

[1]For example, we have accumulated vast evidence in choice under risk (Kahneman, 1979; Holt, 1986; Tversky and Kahneman, 1992; Camerer and Ho, 1994; Holt and Laury, 2002; Gneezy et al., 2006; Andreoni

less is known about the *process* by which individuals make decisions—*how* do individuals decide what to choose? And how does the choice environment affect the choice processes people use? The revealed preference tradition is silent about this, yet studying choice processes is relevant for many reasons. For example, naturally, different choice processes can lead to different *choices*, so an analyst could improve their predictions by understanding the choice process of the decision-maker. Furthermore, understanding the choice process can help inform theories, for example, by distinguishing theories that make similar predictions but rely on different underlying mechanisms. Additionally, from a more meta perspective, we might also think that the process itself is welfare-relevant if, for example, certain processes are more cognitively costly. Finally, it might be the case that different processes call for different types of policies and interventions if, for example, certain processes are more susceptible to bias or manipulation.

In this paper, we introduce an experimental paradigm in which to study choice processes, and we use this paradigm to test the hypothesis that individuals resort to "procedural" decision-making as the complexity of the decision environment increases. Complexity has been extensively discussed as a main driver of the choice process in the literature; indeed, both older (e.g.,Simon 1955) and more recent (e.g., Oprea 2020) literature in economics argues that people dislike or are unable to implement complex decision processes and therefore resort to simpler ones. However, the existing literature provides little insight into the *general characteristics* of decision-making as the choices become more complex. We fill this gap by characterizing and demonstrating that, as decisions become more complex, decision-makers (DMs) resort to more *procedural* choice processes, which we define as being "describable" to another person. That is, if a DM can effectively describe their decision-making process to another person—and if this other person can accurately replicate the DM's choices based on the description of their choice process—then it is more likely that they are using a procedural decision rule.

We test our hypothesis using two experiments where we exogeneously vary complexity and measure the describability of choice processes. Section II describes the design of our main experiment. In this study, decision-makers make choices over *lotteries*, and we vary the complexity of the lotteries across treatments by varying the lotteries' support size (Puri, 2023): Each lottery has two possible outcomes in the Simple2 treatment, three out-

and Sprenger, 2011, 2012; Puri, 2018; Bernheim and Sprenger, 2020; Bernheim et al., 2022; Fudenberg and Puri, 2022; Oprea, 2022; Puri, 2023), intertemporal choice (Thaler, 1981; Loewenstein, 1987; Loewenstein and Prelec, 1992; Angeletos et al., 2001; Bernheim and Rangel, 2004; Della Vigna and Malmendier, 2006), social preferences (Forsythe et al., 1994; Berg et al., 1995; Kagel et al., 1996; Slonim and Roth, 1998; Cameron, 1999; Glaeser et al., 2000; Cooper and Stockman, 2002; Andreoni and Miller, 2002; Brandts and Charness, 2003; Fehr and Schmidt, 2006; Charness and Dufwenberg, 2006; Fisman et al., 2007; Cooper and Dutcher, 2011), and beyond.

comes in the Simple3 treatment, and ten outcomes in the Complex treatment. We elicit the decision-makers' choice process by surprising them at a random point and asking them to send a message to another participant—who we call the "replicator"—who will try to guess the decision-maker's previous choices given the description of their choice process. Both the decision-maker and the replicator have monetary incentives tied to the accuracy of the replication, so decision-makers are incentivized to describe their choice process so that their choices are replicable.

Our main outcome variable relates to the difference in replication rates across treatments. In particular, we want to identify the *causal* effect of the choice process description on the ability to replicate choices. As a result, we compare replication rates when replicators have access to the choice process description to replication rates when replicators do *not* see the description. This isolates the effect of the choice process description on replicators' ability to guess decision-makers' choices. Thus, our main hypothesis is that choice process descriptions increase replication rates *more* as decisions become more complex or as the support size of the lottery increases.

We find strong support for this hypothesis. In our main sample, we find that the choice process descriptions increase replication rates by 2 percentage points in the Simple2 treatment, 5 percentage points in the Simple3 treatment, and 12 percentage points in the Complex treatment.[2] We take this as evidence that decision-makers are using more describable—and therefore more procedural—choice processes in complex decisions. Furthermore, we focus on the subset of decision-makers whose decisions are perfectly replicated by the replicator. We call these "perfectly replicable" decision-makers and conjecture that these are the decision-makers who are most likely to be using procedural choice processes. We find that choice process descriptions increase the share of decision-makers who are perfectly replicable by 4 percentage points in Simple2, 5 percentage points in Simple3, and 12 percentage points in Complex. Thus, we identify both more procedural decisions and more procedural decision-makers as complexity increases.

We then turn to analyze the *choices* that decision-makers make and whether the use of procedures affects these choices. We designed some of our lottery choices to test specific hypotheses. First, we repeated lottery menus, and we find that procedural decision-makers are more likely to choose *consistently* across repeated menus. We also find that procedural decision-makers are less likely to violate dominance. While the impact of procedural decision-making on choice quality likely varies by domain, these results suggest that procedural decision-making can help prevent "mistakes" in some choice environments and,

[2]As noted in Section II.A, our main sample excludes menus with "obvious" choices since these mitigate any treatment differences; that said, our results hold in the full sample.

more generally, suggest that increased use of procedural decision-making affects the actual choices that individuals make. This affirms the usefulness of studying choice processes and establishes that these choice processes are welfare-relevant.

We conjecture that individuals tend toward procedural decision-making as complexity increases because more procedural choice processes are *easier* to implement and thus reduce the cognitive costs of decision-making. We confirm this in secondary evidence: As complexity increases, decision-makers are more likely to self-report using "shortcuts," and decision-makers who report using shortcuts are more replicable. We also find that replicators guess the decisions of perfectly replicable decision-makers significantly faster than they guess the decisions of others, suggesting that the perfectly replicable choice processes are easier to implement.

Finally, we ask what procedures are individuals using. Even if our experimental design is not designed to answer this question, one of the features of our methodology is that the DMs' messages, when leading to accurate replications, give precise individual-level descriptions of their choice process. This allows us to provide very rich and unique evidence on which procedures specific DMs implement. At the same time, summarizing the procedures that DMs use is, of course, very non-trivial: Individuals have heterogeneous preferences and likely use many different procedures. We indeed find that procedures are very heterogeneous and that some of the most simple procedures we can identify (e.g., always choosing the lottery with a higher expected value) do not seem to be driving the procedural response that DMs exhibit in the face of complexity.

We replicate our main results in a second experiment that we describe in Section III. This serves as an important robustness of our main mechanisms and extends our results in a few important directions—in particular, it allows us to test our main hypothesis under a different notion of complexity. In this second experiment, we have individuals make choices between *charities*, and we vary complexity by the *cardinality* of the menu. In the Simple treatment, individuals choose to make a donation to one of *two* charities in each menu, while in the Complex treatment, they choose to make a donation to one of *six* charities in each menu. We replicate our main results in this experiment—choice process descriptions increase replicability rates more in the Complex treatment than in the Simple treatment.

We add to our results on the choice process by showing that decision-makers in different treatments are choosing different charities by estimating a discrete choice model for each treatment that puts significantly different weights on the different charity characteristics. Discrete choice models are precisely structured algorithms, so decisions made with structured algorithmic approaches might be better approximated by these tools. We show that, indeed, our choice model fits procedural decisions better.

We replicate our results in the risk experiment that support the conjecture that individuals tend toward procedural decision-making as complexity increases because more procedural choice processes are *easier* to implement and thus reduce the cognitive costs of decision-making. We confirm this in secondary evidence: As complexity increases, decision-makers are more likely to self-report using "shortcuts," and decision-makers who report using shortcuts are more replicable. We also find that replicators guess the decisions of perfectly replicable decision-makers significantly faster than they guess the decisions of others, suggesting that the perfectly replicable choice processes are easier to implement.

Our results contribute to a few literatures. Most directly, we speak to the literature that has discussed procedural decision-making, starting from at least Simon (1955). Notably, Simon (1976) popularizes "procedural rationality" as behavior that "is the outcome of appropriate deliberation." He directly contrasts this with "substantive rationality"—the typical notion of rationality in economics—which he defines as behavior that "is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints." While Simon's notion of "procedural rationality" certainly carries a normative implication, our current approach is mostly descriptive in nature. That said, there is clearly scope for future work to think about normative questions, given, for example, that we find a connection between procedural decision-making and decision quality and cognitive costs. Since Simon's seminal work, many theoretical contributions have incorporated procedural elements that depart from the traditional Von Neumann and Morgenstern (1944) expected utility framework (e.g., Dubra and Ok 2002; Rubinstein and Salant 2006; Manzini and Mariotti 2007; Salant and Rubinstein 2008; Mandler et al. 2012; Cherepanov et al. 2013; Lleras et al. 2017). For example, in the context of choice under risk, Rubinstein 1988 formalizes a similarity heuristic by which people evaluate two distinct alternatives ignoring the attributes of these alternatives that are "similar" to each other (see Huber et al. 1982; Tversky and Kahneman 1986; Rubinstein 1998 for related empirical evidence), and Gilboa and Schmeidler 1995 suggests a decision rule that uses past performance in similar cases (more recently, Bonder et al. 2023 interact this idea with complexity).[3] Much of this work makes evident the importance of understanding mechanisms to select between theories that have the same predictions (e.g., Mandler et al. 2012 compares the checklist procedure with utility maximization).

Most empirical papers have focused on specific procedures that individuals might be using in a given environment, often trying to infer the use of procedures from choice data (e.g., Payne et al. 1988, Halevy and Feltkamp 2005; see Johnson et al. 2008 for a discussion of

[3]Rubinstein's similarity heuristic is similar to the "cancellation" heuristic (Wu, 1994; Weber and Kirsner, 1997; Wakker, 2022).

the role of process data.). For example, Choi et al. (2006) uses a portfolio choice problem and finds the prevalence of choices consistent with the use of a few heuristics, like trying to guarantee a minimum payoff level. Relatedly, Halevy and Mayraz (2022) study a portfolio choice problem in which, on top of allocating on a case-by-case basis, participants design an investment rule for allocating their budget. Their design allows participants to choose the case-by-case or the investment rule method for a third set of problems, and they find that most participants prefer to use investment rules.[4] An essential difference between our design and the experimental contribution of Halevy and Mayraz (2022) is that we are able to capture the endogenous (i.e., untriggered) use of procedures.[5] Moreover, they focus on distinguishing whether participants behave "as-if" they are maximizing a utility function, or "as-if" they are implementing a decision rule, which would look indistinguishable from choice data only. We contribute to such efforts: By directly eliciting the choice process, our methodology is able to get to the core of this distinction, which is an example of how understanding the choice process can inform theory by helping distinguish those with similar predictions but different mechanisms. This connects us to literature that goes beyond choice data to study the choice process itself, both in psychology and economics (e.g., Reutskaja et al. 2011; see Schulte-Mecklenbeck et al. 2017 for a review). Our experimental methodology also allows future research to directly measure a hypothesized procedure of interest; we discuss this more in Section V.

Ideas very close to our approach are present in the work of Heiner (1983) who not only emphasizes the intuition that "observed regularities of behavior can be fruitfully understood as "behavioral rules" that arise because of uncertainty in distinguishing preferred from less-preferred behavior," but also draws a connection intimately related to how we characterize procedures in this paper, that "[behavioral rules] simplify behavior to less-complex patterns, which are easier for an observer to recognize and predict."[6]

More generally, we join a long literature on bounded rationality that shows the ways in which individuals simplify their decision-making (Simon 1955, 1976; Aumann 1997, 2019). This includes both papers in the tradition of heuristics and biases (Tversky and Kahneman, 1974; Luce, 1978; Kahneman et al., 1982; Rubinstein, 1988; Gigerenzer and Goldstein, 1996;

[4]Relatedly from an experimental design point of view, Romero and Rosokha (2018), Dal Bó and Fréchette (2019) and Cason and Mui (2019) allow participants to design a strategy in the context of an infinitely repeated prisoners' dilemma.

[5]While the interesting results in Halevy and Mayraz (2022) do not show that the take-up of the investment rule increases with menu size, we don't interpret this to be in contrast to our results in terms of the use of procedures in the face of complexity, as participants can very well be using procedures of the type we capture even without choosing the investment rule method.

[6]While Heiner contrasts behavioral rules and their predictability to optimizing behavior, we find evidence that procedures can improve the quality of choice. This evidence is not, however, necessarily inconsistent with Heiner's conceptual framework.

Gigerenzer and Todd, 1999; Benartzi and Thaler, 2001; Gigerenzer and Selten, 2002; Gilboa et al., 2009; Gigerenzer and Gaissmaier, 2011, among many others), as well as papers studying attention allocation (e.g., Gabaix et al. 2006), such as those studying rational inattention (Sims, 2003; Masatlioglu et al., 2012; Caplin and Dean, 2015; Caplin et al., 2019, 2020, 2022; Maćkowiak et al., 2023). The procedural decision-making we identify certainly relates to both heuristics and rational inattention. Interpreting heuristics as decision-making shortcuts, one can consider procedures as defined in this paper (which can also be thought of as "procedural heuristics" to capture the notion of a shortcut, in light of the evidence presented in section II.C.1) as a subset of broadly defined heuristics.[7] Thus, we identify procedures as an alternative categorization of decision-making processes that boundedly rational agents use in the face of complexity, and in section II.C.1 provide suggestive evidence that they operate as a simplification.

We also contribute to both old and very recent literature on complexity and the behavioral responses to complexity (Simon, 1955; Heiner, 1983; Salant, 2011; Oprea, 2020; Salant and Spenkuch, 2022; Mononen, 2022; Oprea, 2022; Puri, 2023; Enke et al., 2023; Banovetz and Oprea, 2023). In particular, Heiner (1983) recognizes complexity, which he understands as uncertainty in distinguishing preferred from less-preferred behavior stemming from the gap between the environment's complexity and the agent's competence, as the basic source of predictable behavior, making complexity a crucial component of how we model behavior. More recently, and perhaps the closest work to the ideas in this paper, Banovetz and Oprea (2023)— henceforth BO—experimentally tests that people dislike implementing complex procedures and that this aversion drives the use of decision rules. Specifically, they focus on "state complexity" as motivated by the automata literature. We see the current paper and BO as complementary work, where the latter focuses on the complexity of the procedure itself (for example, a procedure that requires keeping track of the history of more states is a more mentally laborious procedure) to show that decision-makers exhibit a preference for simpler procedures, while we concentrate on the complexity of the choice problem and show how decision-makers use more describable processes as the complexity of the problem increases.[8] Importantly, while BO find that, in using procedures, decision-makers deviate from optimality, we show evidence that procedures sometimes improve choice quality, which

[7] This means that, for example, one can think of heuristics that *might* be describable, such as satisficing (Caplin et al., 2011), and others that are likely not, such as updating heuristics –representativeness (Tversky and Kahneman, 1974). Importantly, while our characterization allows us to identify procedures as describable choice processes and, hence, only provide evidence about the increase in describable (and replicable) processes in the face of complexity, our hypothesis intuitively extends to non-describable heuristics, as has been extensively discussed in the literature on bounded rationality cited above.

[8] A point of direct contact between both approaches lies in interpreting the describable choice processes that we identify as "simple procedures" in the language used in Banovetz and Oprea (2023).

emphasizes the rich and varied consequences that the use of procedures has on choices and well-being.[9]

In addition, our specific experiments contribute to the literatures on risk and choice overload. There is a discussion in the literature on choice under risk of whether modern experiments capture risk *preferences* or instead capture heuristic responses to the complexity of the experiment (Bernheim and Sprenger, 2020; Wakker, 2022). Our results do not draw conclusions on this, but we provide tools with which future research can study this question more directly.[10] Finally, our charity experiment implements complexity through choice overload; this literature typically studies the status quo bias as a response to choice overload complexity, and we identify procedural decision-making as an additional response (Iyengar and Kamenica, 2006; Scheibehenne et al., 2010; Chernev et al., 2015; Dean and Neligh, 2017; Abaluck and Gruber, 2023).

Broadly speaking, we believe our experiment opens two avenues for future work. First, future work can understand more about procedural decision-making and the types of complex environments that trigger this response. It is likely that individuals use procedures to make many of their complex decisions, and even decisions that are very high stakes. For example, there is evidence that straight-ticket voting increases with ballot length (Augenblick and Nicholson, 2016), which can be thought of as a procedural response to complexity. It would be very valuable for future work to identify such procedures and their consequences on welfare and other outcomes. Further on this line, our findings allow for procedural reinterpretations of existing findings. For example, identifying which known heuristics are describable versus ones that are not can be useful in understanding their underlying cognitive mechanisms and can assist in interventions that improve decision-making.

Second, future work can use our experimental methodology to understand choice processes beyond the use of procedural decision-making. Seminal work in psychology by Nisbett and Wilson (1977) (henceforth, NW) reviews evidence in support of previous ideas in the literature advocating that individuals have "no direct access to higher order mental processes such as those involved in evaluation, judgment, problem solving, and the initiation of behavior," and describe cases in which subjective reports of cognitive processes prove to be accurate as the "incidentally correct employment of a priori causal theories."[11] We see our approach as fundamentally distinct: Much of the work reviewed in NW asks participants

[9]Indeed, Gigerenzer and Todd (1999) characterize heuristics as *efficient* cognitive processes. More generally, Gigerenzer's work emphasizes how heuristics can improve choice quality.

[10]Importantly, our design includes lotteries of complexity levels most used in this discussion, with 2 and 3 outcomes, as well as more complex lotteries which serve as benchmarks for the identification of choice processes.

[11]Under this lens, our results could be interpreted as showing that individuals' a priori causal theories become more procedural as complexity increases. We do not share this interpretation.

"why" they did what they did, in contrast to our approach of asking "how" their choices came to be with the goal of external replication.[12] Introspectively, this distinction clearly calls for different responses: we are interested in eliciting the choice process, while much of the evidence in the literature cited in NW seems to elicit the motives behind the choice. [13] Our methodology provides a general way to not only *incentivize* choice process descriptions but also assess their accuracy via replication. These descriptions can then be further analyzed using modern experimental and other techniques. We return to both of these directions for future work in our Conclusion.

II. Experiment 1: Risk

We designed two experiments to test our main hypothesis that decision-making becomes more describable and replicable as decisions get more complex. While it is very likely that individuals use procedures to make decisions outside of the lab, a controlled experiment allows us to vary complexity exogenously and allows us to carefully measure choice processes. We conduct our first experiment in a well-studied and economically relevant domain—choice under risk—and we conduct our second experiment in a more naturalistic environment—donations to charities—to validate our findings.

II.A. Risk Experiment Design

Our first experiment compares participants' choice processes when choosing from binary menus of lotteries with varying numbers of outcomes. We elicit participants' choice process by asking them to send a message to another participant who will try to guess their previous choices; both participants are incentivized by the accuracy of the replication.

Our experimental design involves two main types of participants: "decision-makers" and "replicators." We study the choice processes that decision-makers use and use replicators to measure and incentivize the elicitation of the choice process. We first describe the structure of the decision-makers' study and then describe the replicators' study. Screenshots of the instructions for both decision-makers and replicators are included in the Appendix.

[12]In our case, this would mean asking participants "Why did you pick that lottery/charity?" and trying to find –or explicitly cue– references to the role of complexity as manipulated in our study. We do not think this is an enlightening approach for our purposes.

[13]Much of the evidence that drives the paper's conclusions stems from the failures of individuals to mention, in their reports, an element that the researchers manipulate and believe to cause behavior (e.g., the effect of some stimulus the researcher imposes). This dissonance is not only not insufficient to disregard individuals' ability to access their choice process, but it can even be interpreted as a symptom of the researcher's failure to identify the causal mechanism. Smith and Miller (1978) and White (1980) criticize NW and claim their viewpoint is not falsifiable. See Berger et al. (2016) for a more recent discussion on NW and related claims.

II.A.1. Decision-Makers Study

To study the use of procedures, we give decision-makers (DMs) a sequence of 25 incentivized binary choices between lotteries, presented in random order.[14] We present the lotteries in a table with the outcomes and associated probabilities displayed in random order. DMs know that one randomly picked choice from a randomly picked participant will be selected and the outcome of the chosen lottery will be paid to them. Thus, in making their decisions, decision-makers are incentivized to choose the lottery that they most prefer, just as in standard experiments studying choice under risk.

To study how complexity affects the choice process, we randomly assign participants to one of three treatments that the number of outcomes in the lotteries. The support size of a lottery is a natural dimension of complexity and has been discussed in the literature of choice under risk (Puri 2023).

Simple2 Treatment: All lotteries have two outcomes.

Simple3 Treatment: All lotteries have three outcomes.

Complex Treatment: All lotteries have ten outcomes.

Alongside the lottery outcomes and probabilities, DMs in all treatments see five buttons that, when clicked, show one of the following pieces of additional information about the lotteries: expected value, variance, minimum payment, maximum payment, and the probability with which the lottery pays the maximum payment.[15] We selected these pieces of information as they were the most frequently mentioned in our pilot data, relate to procedural discussions in the risk literature, and seem natural to enter into simple procedures DMs could be using. DMs can only see the information displayed by one button at a time, and they see that information for both lotteries in the menu, making comparisons across lotteries very easy. Figure I shows an example screenshot of a choice in the Complex treatment where the "maximum payment" button has been clicked.

This additional information is a feature that allows us to reduce noise in identifying procedures in a few ways. First, DM's know that replicators will also have access to this information, which gives decision-makers a common language to communicate their choice process to the replicators. Second, displaying this information in buttons also enables us to collect additional non-choice data in the form of whether the DM acquired this information.[16] We

[14]DMs knew that they would face 25 choices, which we felt would make the gains from developing a procedure larger and would give DMs experience to come up with procedures over time.

[15]We explain the meaning of these five pieces of information to subjects using common language; see Appendix II.A for the full text.

[16]This is one of the reasons why we displayed the lottery outcomes in random order; in doing so, we make it more valuable to use, e.g., the button that display's the lottery's maximum payment.

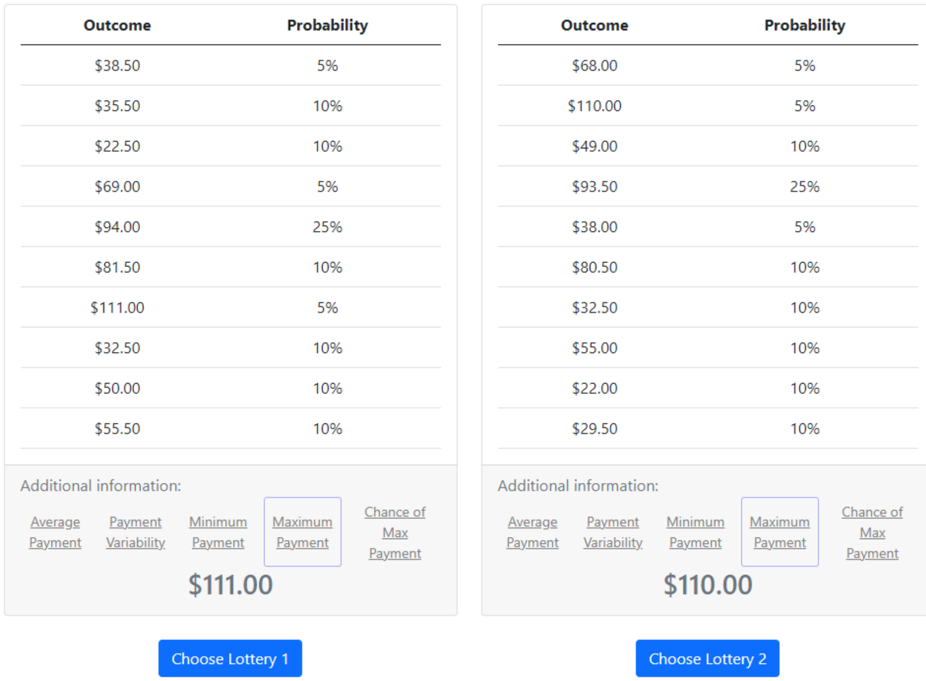


Figure I: Example of lottery menu in the Complex treatment

can then correlate DM choices and messages with the use of the buttons to further validate our measures of choice processes.

The 25 menus that DMs see consist of four types of questions: standard lotteries, lotteries related by dominance, mean-preserving spreads, and repeated choices. 17 of the 25 menus are "standard" lottery choices. We construct these 17 menus by, within each treatment, randomly generating lotteries with the requisite number of outcomes and randomly matching these lotteries together to create binary menus. We randomly re-matched lotteries until the distribution of the differences of the following moments were similar across treatments: expected value, variance, minimum outcome, maximum outcome, and chance of maximum outcome.[17] Appendix section I.C shows these distributions. Four menus consisted of lotteries related by dominance, including both First Order Stochastic Dominance (FOSD) and statewise dominance[18] Two menus consisted of a lottery and a mean-preserving spread of the lottery. Finally, we repeated two randomly selected menus.[19]

We included dominance and mean-preserving spread menus to test whether procedural DMs would be more or less likely to make "mistakes"—dominance violations are commonly

[17]We construct lotteries in this way because significant differences in these moments between the lotteries in the menu across treatments would introduce confounds when interpreting the treatments' differences as the causal effect of complexity.

[18]In Simple2, these notions coincide.

[19]We do not repeat the dominance menus given that we already had four potential dominance violations.

thought of as mistakes in risky choice, and while choosing a mean-preserving spread is not an obvious mistake, most experimental participants display small stakes risk aversion and do not choose mean-preserving spreads. Thus, we can test whether the use of procedures leads to more or less dominance violations and choice of mean-preserving spreads. We included repeat menus to test whether different choice processes affect the consistency of choices, i.e., whether procedural DMs are more or less likely to choose the same option in a repeated menu.

At a random round between rounds 10 and 25, DMs are surprised with the *Message Task*. The Message Task elicits the DMs choice process by asking them to describe to another participant *how* they made their last five decisions. Specifically, we present participants with the following taks:

- Message Task Prompt: *Please write a message to another participant describing how you made your last five decisions.*

We incentivize the Message Task by telling DMs that another participant will see this description of the choice process and try to guess their last five choices and that they may earn a $5 bonus payment if the other participant is accurate in one randomly picked guess. They also know the other participant will get the same bonus if the guess is accurate. Thus, DMs have an incentive to describe their decision-making process as accurately as possible.

To identify procedural decision-making, we want DMs to describe their decision-making *process* rather than describing their *choices*. We incentivize this in a few ways. First, DMs are not able to view past decisions or menus when writing their message. Second, DMs know that the other participant will see their last five decisions in random order and with the lotteries relabeled: What was lottery one in decision four for the DM need not be the same for the other participant. Finally, the message elicitation comes as a surprise to the DMs, so DMs have no incentive to attempt to remember their decisions or change their decision-making process while they are making choices.

After submitting their message, we elicit DMs beliefs about how many decisions, out of five, the replicator will accurately guess. We incentivize this elicitation by randomly picking a DM and giving them an additional $10 if their guess is accurate.

After DMs make all 25 decisions, we ask a few unincentivized survey questions. These included questions such as whether they developed a shortcut to pick lotteries and how easy it was for them to decide which lottery to use. We include the list of all questions in the Appendix.

II.A.2. Replicators' Study

We want to test our hypothesis that the decision-making process becomes more describable as decisions become more complex and, therefore, that the DM's message better describes their choices in complex decisions. We identify choice processes that are easy to describe as those that are easy to *replicate* based on the description of the choice process. However, comparing replicability rates across treatments at face value introduces a confound—it could be the case that some decisions are simply easy to replicate for reasons that have nothing to do with the choice process. For example, one lottery in the menu might be very salient to both DMs *and* replicators, leading to very high replicability rates even though the choice process might not be describable. Furthermore, since the lotteries and menus vary across treatments, it could be the case that this is more likely in some treatments than others, muddying the comparison of replicability across treatments. To solve this issue in identification, we isolate the causal effect of the choice process description on replicability rates by comparing replication rates *with* and *without* the description.

Specifically, we recruit a new group of participants that we call "replicators" (R). We randomly assign replicators into one of two conditions that vary *only* in whether replicators have access to the DMs description of their choice process:

Message Condition: Replicators see the DMs description of their choice process and try to guess their choices

No Message Condition: Replicators try to guess the DMs choices without seeing their description of their choice process

We tell replicators in both conditions about the DM study, and we show replicators the full set of DMs' instructions, including the DMs' instructions about lotteries in general, the additional information on the lotteries, and the DMs' instructions during the Message Task for replicators in the Message Condition. Replicators answered two understanding questions about lotteries in general and one question about their own incentives in the replication task. Those in the Message Condition also answered one understanding question about the DM's incentives when writing the message.

We randomly match each replicator to three DMs from the same treatment, so each replicator makes a total of 15 guesses. For each DM that they are matched to, replicators see the five menus that the DM saw before they faced the Message Task and are asked to select the lottery they think the other participant chose in that decision. The lotteries are displayed to replicators *exactly* as they are displayed to DMs—including the additional information in buttons—but the order of the lotteries on the screen is randomized. Replicators in the Message Condition see, above the two lotteries, the message that the DM wrote describing their choice process.[20] We incentivize the replication task in both conditions by telling replicators that they may earn a $5 bonus payment if a randomly picked guess is accurate so that they are incentivized to guess as best they can.

After going through the five replication exercises for a given DM, we elicit replicators' beliefs about how many decisions they accurately guessed. We incentivize this elicitation by randomly picking a replicator and giving them an additional $10 if their beliefs are accurate. In the Message Condition, we also ask them unincentivized questions about whether they found the message comprehensible, whether they found it easy or hard to guess based on the message, and whether the message felt like a step-by-step (or single-step) process. In the No Message Condition, after replicating all three DMs, we ask replicators whether they guessed the other participant's choice by selecting what they would have chosen themselves.

II.A.3. Identification

As described above, we will test our main hypothesis—that decision-making becomes more describable as decisions get more complex—by comparing the *difference* in replication rates between the *NoMessage* and *Message* conditions across treatments. However, an additional identification issue remains. It could be the case that the replication rates in the *NoMessage* condition itself differ across treatments. This could occur if procedures are more salient to replicators in one treatment compared to another, for example, and muddies the interpretation of an observed—or lack of observed—difference across treatments.

Since replication rates in the *NoMessage* condition are largely driven by how "obvious" the DM's chosen lottery is, we attempt to make the treatment comparison cleanest by focusing on the menus that are the least obvious. These are menus that maximize disagreement among DMs' choices. Note, we could have designed the experiment *only* to have non-obvious menus, but we did not do so for a few reasons. First, we wanted to include questions related to dominance and these differ in obviousness across treatments. Second, a prior, it is not clear whether decision processes would be affected by facing only non-obvious menus. Since decision-making outside of the laboratory often involves both easy and difficult choices, we decided to include both types of menus. Thus, we randomly generated lotteries and we identify obviousness ex-post through choice probabilities. Specifically, for each treatment and for each menu with that treatment, we identify the most frequently chosen lottery from that menu and we calculate the percentage of DMs who chose this lottery. This is a measure of "obviousness' of the menu that ranges from 50% to 100%. Then, aggregating across all menus separately for each treatment, we identify the menus that fall *below* the median

[20] In both conditions, replicators have access to their instructions and to the descriptions of the additional pieces of information while they are making their guesses.

obviousness level in that treatment.[21]

Our main results will focus on these menus, as they are most likely to equalize replication rates in the *NoMessage* conditions, and they give the most room to identify differences across treatments. Given that our experiment is a proof-of-concept to identify procedural decision-making, it is natural to focus on decisions where we can test our hypothesis most cleanly. However, our results are qualitatively robust to considering all menus, and we present this in the Appendix and note specific references throughout the text. All of the results that follow consider only menus that fall below the median obviousness level unless specifically noted otherwise.

II.A.4. Implementation and Recruitment Details

We recruited all participants on Prolific, an online platform frequently used for research studies.[22]

We recruited 1508 participants for the DMs study across all three treatments. Each participant received a $3 completion payment and took around 17 minutes to complete the study on average. We recruited 963 participants for the replicator study across both conditions.[23] Each participant also received a $3 completion payment and took around 15 minutes to complete the study.

In all study versions, participants receive ample instructions and are required to correctly answer understanding questions before proceeding to the main parts of our study. Rather than excluding participants, they are given as many times as needed to correctly answer the understanding questions. For full experimental instructions of all study versions that we run, see the supplemental Appendix.

II.B. Experiment 1 Results

Figure II shows replicator accuracy rates in the *Risk Experiment*, separated by treatment (*Simple2*, *Simple3*, *Complex*) and by whether the replicator had access to the DMs message (*NoMessage*, *Message*). Replication rates in the *NoMessage* conditions do not differ

---

[21] We pre-registered the analysis of menus by "obviousness" based on the DM choice probabilities in AEARCTR-0010977.

[22] We recruited participants for the DMs study on July 27^{st} of 2023, and those for the replicators study on July 31^{st} and August 1^{st} of 2023. In order to qualify for our study, participants were required to be located in the USA, be fluent in English, and have a minimum of 100 prior submissions on Prolific, with an approval rate of at least 98%. The experiment was implemented using the oTree platform (Chen et al., 2016).

[23] Because we match three DMs to each replicator, we require 500 replicators for each of the *Message* and *NoMessage* conditions for a total of 1000 replicators. We matched 1000 replicators to the DMs and ended up with fewer data points because of participant attrition.

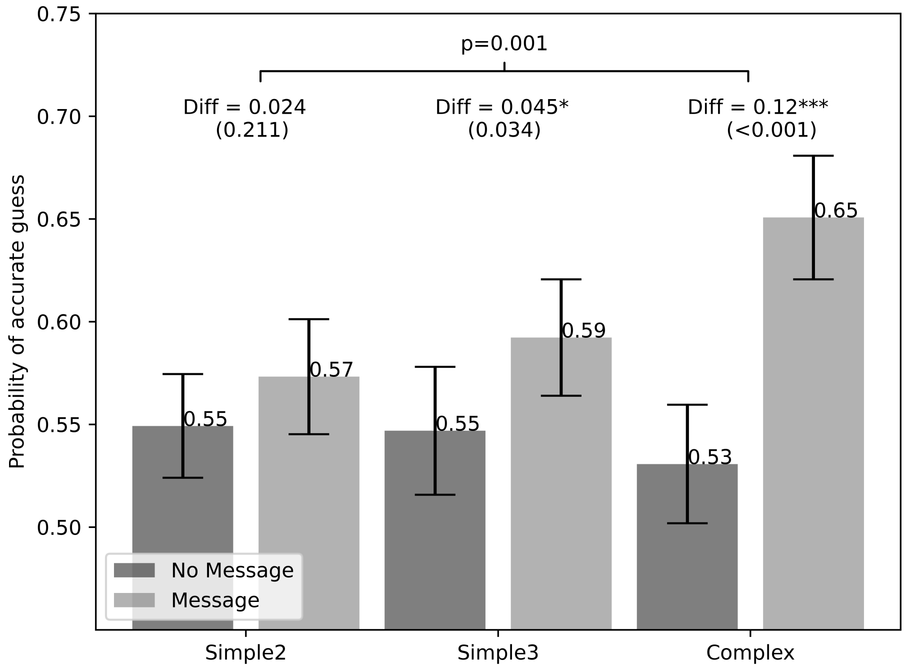


Figure II: Replication accuracy by treatment and condition.
Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments and conditions. The sample considered is the sample of non-obvious menus, as described in Section II.A; see Figure VII in the Appendix for the full sample without this restriction. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

significantly across treatments (55%, 55%, 53%, $p = 0.607$). Our main result compares the *difference* in replication rates between the *NoMessage* and *Message* conditions across treatments, a measure that we refer to as "message gain," or the increase in replication rates when replicators have access to the message. We find a message gain of 2.4 pp in Simple2, 4.5 pp in Simple3, and 12.0 pp in Complex. Interpreted in light of our main hypothesis, we find that messages are *more* effective at helping replicators accurately guess DMs' choices in more complex decisions, suggesting that DMs are using more describable—and therefore more procedural—decision-making processes as complexity increases.

We find that this holds in aggregate—as shown above in Figure II—but it also holds across the distribution of individual decision-makers (see Figure XII in the Appendix for the cumulative distributions of message gain by treatment). We focus on a subset of DMs for whom our hypothesis is particularly salient: DMs for whom the replicator guessed accurately in *all* choices. We refer to these as "perfectly replicable" DMs, and these are the DMs we might think are the "most" procedural in the *Message* condition since the replicator can perfectly guess all of their choices when they have access to the DMs choice process.

Figure III shows the share of perfectly replicable DMs across treatments. We find that

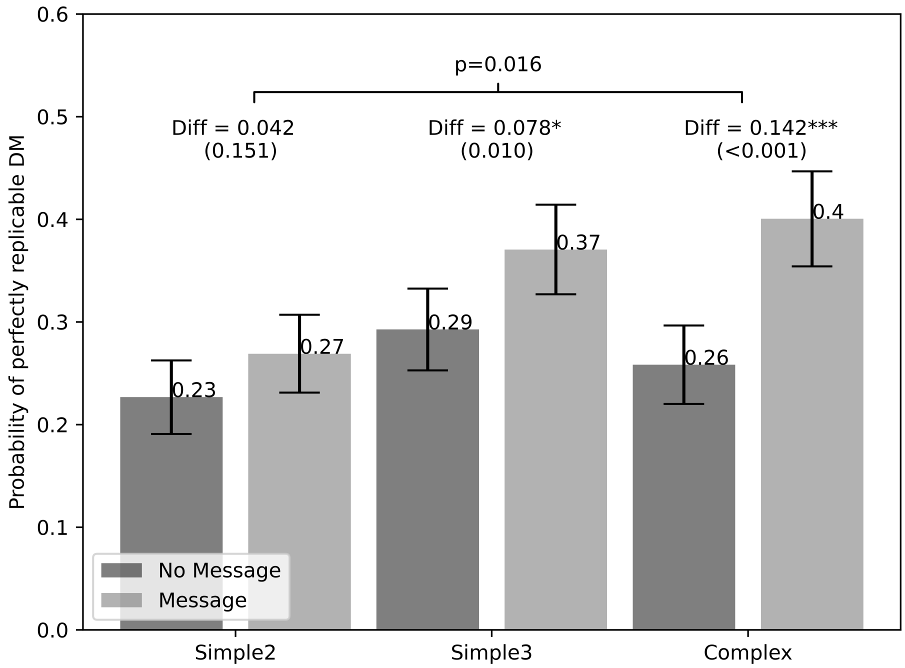


Figure III: Share of perfectly replicable DMs by treatment and condition.
Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments and conditions. The sample considered is the sample of non-obvious menus, as described in Section II.A; see Figure VIII in the Appendix for the full sample without this restriction. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

the message increases the share of DMs that are perfectly replicable by 4.2 pp in Simple2, 7.8 pp in Simple3, and 14.2 pp in Complex. Thus, consistent with our results above, we find that the share of perfectly replicable DMs increases significantly more with the introduction of the message as decisions become more complex (the test of difference in differences yields $p = 0.016$ between Simple2 and Complex).

Finally, in the post-survey questionnaire, we asked participants the following question: "Would you say that you developed a rule or procedure to pick a lottery?" Panel A of Figure IV shows DM responses by treatment. We find that DMs in the Complex treatment are significantly more likely to self-report using a rule or procedure ($p < 0.001$). Panel B of Figure IV shows the share of perfectly replicable based on DMs' self-reported use of a rule or procedure. We find that DMs who say that they used a rule or procedure are significantly more likely to be perfectly replicable than those who say they did not use a rule or procedure ($p = 0.009$). That is, DMs who say they used a procedure are better able to describe their decision-making process to their replicator in a way that enables the replicator to guess their decisions. This further validates that our message gain measure does, indeed, capture procedural decision-making and shows that procedural decision-making increases with complexity.

Panel A. Responses by Treatment **Panel B.** Perfectly Replicable by Response

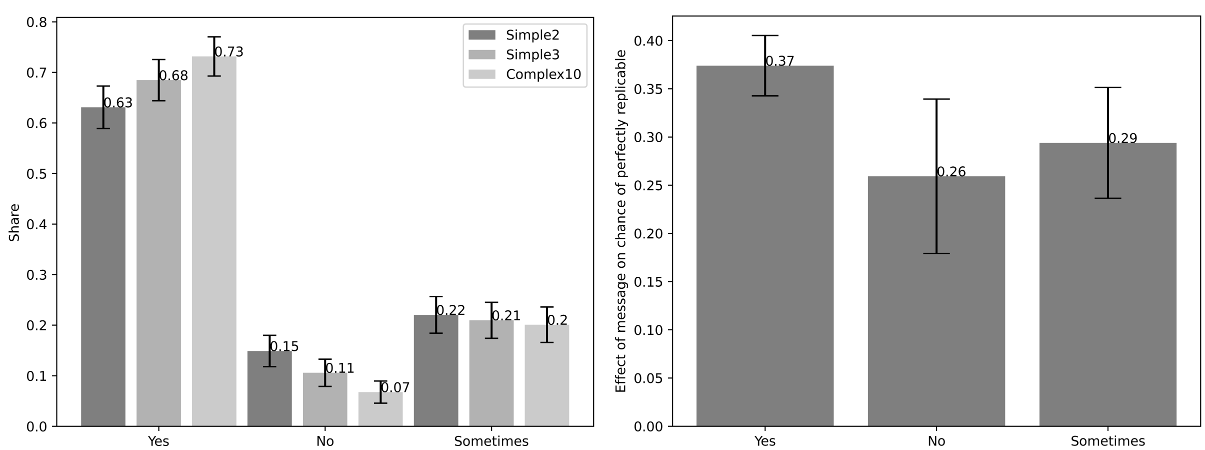


Figure IV: DMs' responses to the question "Would you say that you developed a rule or procedure to pick a lottery?" (panel A), and the share of DMs that are perfectly replicable by their response to this question (panel B).

Note: Bars in panel A show the share of DMs that give each response in the x-axis, by treatment. Bars in panel B show the share of DMs that are perfectly replicable for each response in the x-axis. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

Table I shows the robustness of our main result: using all data collected, controlling for the role of menu obviousness and round, messages in the complex treatment increase accuracy in replication 6 times more than those in the simple2 treatment (0.072 versus

0.011).[24]

Result 1. *Decision-making becomes more describable as decisions get more complex, both on average and at an individual level. We interpret this as decision-makers using more procedural decision-making processes.*

II.C. Does procedural decision-making affect the chosen lotteries?

Having established that decision-making becomes more procedural as choice problems become more complex, a natural following question is whether procedural decision-making affects the actual *choices* DMs make. We have a few ways of looking at this.

We designed specific menus to test whether procedural decision-making affects the choices that DMs make. Specifically, we repeated menus to test choice consistency and included both dominance and mean-preserving-spread menus to test choice quality.[25]

We find that DMs are equally likely to make consistent choices in repeated menus as complexity increases. Specifically, across all repeated menus, the share of DMs who are inconsistent in at least one repeated decision is 43.3%, 39.8%, and 39.0% in Simple2, Simple3, and Complex, respectively ($p > 0.1$ in all pairwise comparisons). However, we find that procedural DMs are more consistent than their non-procedural counterparts. Specifically, aggregating across treatments, 26% of perfectly-replicable DMs (defined above) choose inconsistently in at least one repeated choice, while a significantly larger 43% of non-perfectly-replicable DMs choose inconsistently at least once ($p < 0.001$).[26] Thus, we find consistent evidence that procedural DMing leads to more consistent choices. Note, we do *not* claim that higher consistency is a measure of increased decision "quality" in any way. Instead, these measures support the idea that procedural DMs develop and then implement a decision rule and that implementing an established rule is likely to lead to consistent choices.[27]

Turning to our choice measures that more directly relate to decision quality, we look at our menus involving dominance and mean-preserving spreads. Consistent with evidence from Puri (2023), we find that individuals are *more* likely to violate dominance as the complexity

[24]Our results are also robust to substituting our replicator participants for GPT-4. We feed all messages to the model and ask: "Does it seem like the respondent was using a procedure in choosing a lottery?" and "How procedural does the choice process the participant describes feel from 1 to 5, where 5 is the most procedural and 1 is the least?" For both measures, the model identifies significantly more procedural decision-making among messages from the complex treatment relative to the simple treatment ($p < 0.001$).

[25]For the results on consistency, dominance, and mean-preserving spreads, we use all of the relevant menus without restricting based on "obviousness" since these menus are often obvious.

[26]The gap holds within each treatment but is smaller in the complex treatment, where it holds yielding $p = 0.056$. See details in appendix figure XIII.

[27]Halevy and Mayraz (2022) discusses how consistent choices can be represented by a utility function making "as-if" utility maximizers equivalent to "as-if" implementers of a decision rule, absent non-choice data.

Simple2 Dummy	0.0178
	(0.039)
Simple3 Dummy	-0.0406
	(0.016)
Complex Dummy	-0.0189
	(0.015)
Message Dummy	0.0772
	(0.076)
Message * Simple3	0.0379
	(0.098)
Message *Complex	0.3168
	(0.102)
Menu Obviousness	0.9320
	(0.044)
Obviousness * Simple3	-0.0274
	(0.053)
Obviousness * Complex	-0.2086
	(0.054)
Obviousness * Message	-0.1329
	(0.079)
Obviousness * Message * Complex	-0.2086
	(0.054)
Obviousness * Message * Simple3	-0.0274
	(0.053)
Surprise Round	-0.0021
	(0.001)
Round * Simple3	0.0008
	(0.002)
Round * Complex	0.0011
	(0.002)
Round * Message	0.0016
	(0.003)
Round * Message * Simple3	0.0008
	(0.002)
Round * Message * Complex	0.0011
	(0.002)

Table I: Effect on accuracy of replication

Notes: This report results from an OLS regression. The dependent variable is a binary indicator for the accuracy of the guess. The dummy for simple2 is estimated as the constant in the regression (omitted category). We report standard errors in parentheses; standard errors are clustered at the replicator-level. See appendix table III for details.

of the choice problem increases— 19%, 35%, and 48% of DMs violate dominance in Simple2, Simple3, and Complex, respectively ($p < 0.001$ in all pairwise comparisons). However, we find that procedural DMs are significantly *less* likely to violate dominance. Splitting by DMs that are perfectly replicable versus those that are not and aggregating across treatments, we find that perfectly replicable DMs are almost half as likely to violate dominance relative to not perfectly replicable ones (20% compared to 38%, $p < 0.001$). While this is an example of procedural decision-making *preventing* a type of "mistake," this is unlikely always to be the case. Given that procedures are simplified choice processes, there are many environments where procedures are likely to lead to lower-quality decisions.

We don't find any consistent evidence on procedural decision-making affecting the choice of mean-preserving spreads across treatments. While there is suggestive evidence that individuals are more likely to choose a mean-preserving spread as complexity increases—they choose the spreads 38%, 63%, and 48% of the times in Simple2, Simple3, and Complex, respectively—we find no correlation between choice of mean-preserving spread and perfectly replicable decision-makers (the perfectly replicable choose a mean-preserving spread 71% of the time, and the not perfectly replicable 68%; $p = 0.423$).

Taken together, we find evidence that procedural decision-making changes the choices that individuals make. This could happen through the channel of information acquisition or through the use of procedures that can even help prevent certain "mistakes" and non-standard choice patterns documented in the literature. Given that procedural decision-making changes choices, it, therefore, would change the inference an analyst would make from these choices.

Result 2. *Procedural decision-makers acquire more additional information, are more likely to choose consistently across repeated menus, and are less likely to violate dominance as complexity increases. These results suggest that procedural decision-making changes the choices that individuals make.*

II.C.1. Are procedures simplifying the choice process?

We hypothesize that individuals use procedural decision processes because this type of decision-making is "easier" to implement, which is why these procedural DM processes are more common in complex decisions. We test this hypothesis in a few ways.

First, as a proxy for how difficult a choice process is to implement, we can look at how long it takes for replicators to guess DMs' choices. The idea is that replicator response times give us a measure of how difficult the choice process—taken as a "decision rule"—is to implement. Thus, the hypothesis is that simple choice processes are easy to implement and, therefore, will result in choices that are replicated very *quickly*. We find evidence consistent with this: Perfectly replicable DMs (as defined above) are replicated in 16.5 seconds on average across treatments, while non-perfectly-replicable DMs take significantly longer to replicate at 18.3 seconds (a test for the difference yields $p = 0.017$).[28]

Second, in our follow-up survey, we simply asked DMs whether they used a "shortcut" to make their decisions.[29] We find that DMs are significantly more likely to self-report using a shortcut in the Complex treatment—specifically, 18% of DMs in Simple2, 30% of DMs in Simple3, and 37% of DMs in Complex answered "yes" when asked if they used a shortcut ($p < 0.001$ for all comparisons to Simple2, and $p = 0.015$ for the difference between Simple3 and Complex). As validation of our self-reported measure, we find that DMs who self-report using a shortcut are, indeed, more procedural according to our measure. Specifically, aggregating across treatments, we find that 47% of DMs who reported using a shortcut are perfectly replicable, compared to only 29% of those who reported not using a shortcut or using it sometimes ($p < 0.001$).[30] Thus, we find that DMs who reported using a shortcut are more procedural according to their associated replication rates.

As a final piece of suggestive evidence, we asked DMs in the follow-up survey how "easy" they found it to choose their preferred lottery.[31] We find no significant differences in DMs responses (see Appendix Figure XIV). In all three treatments, about 26% of individuals found it very easy, 62% somewhat easy, 11% somewhat difficult, and less than 0.5% very difficult. Thus, even though DMs in Complex were making decisions with 3–5 times more outcomes per lottery, they did not report that the decisions were significantly harder to make. This suggests that the use of simpler choice processes mitigated the cognitive cost of difficult decisions. As validation of this self-reported measure, we find that DMs who self-report finding it "very easy" to pick a lottery are, indeed, more likely to be procedural according to our measure. We show these results in Figure XV. Specifically, aggregating across treatments, we find that 39.3% of DMs who reported it was very easy are perfectly replicable, compared to 34.2% for those who reported it was somewhat easy ($p = 0.097$), and 26.6% for those who reported it was somewhat difficult ($p = 0.007$).[32] Thus, we find that DMs who reported finding it easier to pick are more likely to be procedural according to

[28]We find that DM response times are significantly longer in the Complex treatment overall; see Figure XVI in the Appendix for average response times by round in each treatment. Interestingly, *perfectly replicable* decision-maker response times do not differ across treatments; see Figure XVII in the Appendix. We believe replicator response times give a clearer picture of the difficulty of implementing the choice process since DM response times could be conflated by, e.g., the time it takes to *develop* the procedure, among many other things.

[29]Specifically, we phrased the question: "Would you say that you developed a shortcut to pick a lottery?"

[30]This result holds within all treatments.

[31]Specifically, we phrased the question: "How easy was it for you to decide which lottery to choose?"

[32]We also had an option for DMs to report that they found it very difficult. Only 6 DMs picked that option, and none of them is perfectly replicable.

their associated replication rates.

While no one piece of evidence is conclusive, taken together, we find evidence consistent with procedural decision-making as a simplification of the choice process.

Result 3. *Secondary evidence is consistent with the interpretation of procedural decision-making as a choice process simplification.*

II.C.2. What procedures are people using?

We have documented robust evidence that individuals use more procedural decision-making processes as complexity increases, that these procedures help simplify the decision-making process, and that this type of decision-making results in different chosen outcomes. Thus, a final natural question is exactly *what procedures* are individuals using. Even if our experimental design is not designed to answer this question, one of the features of our methodology is that the DMs' messages, when leading to accurate replications, give precise individual-level descriptions of their choice process. This allows us to provide very rich and unique evidence on which procedures specific DMs implement. At the same time, summarizing the procedures that DMs use is, of course, very non-trivial: Individuals have heterogeneous preferences and likely use many different procedures; message data is extremely rich, which can make it hard to systematically analyze; and the universe of possible procedures is extremely vast, which makes aggregation hard. However, we take a first step toward answering this question in a few simple ways that we present below. Overall, we find that procedures are very heterogeneous and that some of the most simple procedures we can identify do not seem to be driving the procedural response that DMs exhibit in the face of complexity.

As a first step, we try to shed light on the procedures DMs use by exploring the *information* that DMs attend to in making their decisions. Our experimental design captures non-choice data that allows us to assess whether DMs acquire different information before making their choices in the form of the buttons that DMs click on. For all of our buttons, DMs are significantly more likely to acquire the information as complexity increases; see figures XXI and XXII in the appendix. This does not necessarily mean that this information necessarily enters the DMs' choice process, or that this is the only information that enters the process—for example, it's quite likely that individuals could be using some of this information without needing to click on it in the Simple treatments—but nevertheless, it suggests that DMs might be using different information to make their decisions as complexity increases.

The simplest procedure we might think DMs use is to maximize a single dimension corresponding to one of the provided buttons, e.g., always choosing the lottery with a higher expected value. We can look for evidence of this in a few ways.

First, we identify DMs who—in the five choices prior to the message task—clicked the same single button in each of the five choices. That is, we identify decision-makers who, e.g., clicked on average payment and only average payment in all five decisions. We call these "single-button" DMs. Separating DMs by those that are perfectly replicable versus those that are not, we find that 25% of perfectly replicable DMs are single-button while only 14% of non-perfectly-replicable DMs are single-button ($p < 0.001$). While this result confirms the intuition that these simple procedures are, naturally, very replicable, our findings suggest this type of procedure is *not* driving the replicability gap across treatments. The share of DMs that are single-button is indistinguishable across our simplest and most complex treatments, at 15.5 and 15.8%, respectively ($p = 0.899$).[33]

Along these same lines, we can look at the *choice data* to identify individuals who are consistent with maximizing a single dimension. Similar to above, we look at the five choices prior to the message task and see whether the DMs' decisions from these menus can be rationalized by maximizing/minimizing one of the dimensions corresponding to the buttons (maximizing EV, minimizing variance, maximizing max payment, maximizing min payment, or maximizing the chance of the highest prize). This exercise leads us to similar conclusions as that of analyzing single-button DMs: We confirm the intuition that these simple procedures are very replicable, and our findings suggest this type of procedure is not driving the replicability gap across treatments. Specifically, separating DMs by those that are perfectly replicable versus those that are not, we find that 51% of perfectly replicable DMs are consistent with maximizing one dimension while only 36% of non-perfectly-replicable DMs are ($p < 0.001$). However, we find that 44% of DMs in Simple2 can be rationalized by maximizing a single dimension, compared to 43% in Simple3 and 35% in Complex (comparing Complex with Simple2 and Simple3 yields p-values of 0.009 and 0.032, respectively. The comparison of Simple2 and Simple3 yields a p-value of 0.702).

Thus, we find that the simple procedures we might have ex-ante hypotheses DMs use are, indeed, very replicable. However, it is not the case that DMs use these simple procedures more as decisions become more complex, and, therefore, it is not the case that these simple procedures are what is driving the increase in procedural decision-making in general.

Finally, we directly analyze our message data to uncover insights into the type of procedure DMs are using. Interestingly, we do find that messages get longer as complexity increases, yet the length of the message seems unrelated to replicability; see appendix figures X, IX and XI for details.[34]

[33] Our simple treatment happens to host significantly more single-button DMs at 21.9%. These levels could be affected by the idiosyncrasy of the lotteries DMs face in each treatment.

[34] The average message is 169, 185, and 198 characters long in the simple2, simple3, and complex treat-

We hypothesize that many procedures follow a more "step-by-step" or algorithmic structure.[35] We identify this by looking for evidence of this step-by-step logic in the message content. Specifically, we code for language that is indicative of step-by-step messages such as "first,", "one" and "then," and we call these *step-by-step messages*.

We find that significantly more messages in Complex are classified as step-by-step (33.7% vs 37.9%, and 43.0% for Simple2, Simple3, and Complex, respectively. Comparison of Complex to Simple2, Simple3 yields p-values of 0.005 and 0.137, respectively.), suggesting that decision processes in complex decisions become more structured and algorithmic (appendix figure XVIII plots these percentages across treatments).[36] Furthermore, after replicators guessed the five decisions for a DM, we asked them whether the message felt like a step-by-step (or single-step) process. Replicators matched to DMs from the Complex treatment were significantly more likely to indicate that the message felt step-by-step, compared to replicators matched to DMs from the Simple2 treatment (see Figure XIX in the Appendix).

While we do not find that perfectly replicable DMs are more likely to use step-by-step language as identified by our coding of messages, we do find that perfectly replicable DMs are more likely to have replicators who said the message felt step-by-step.[37] Our message classification is very simple and likely does not capture all step-by-step messages, but results suggest that replicators who felt that they were implementing a step-by-step algorithm are more likely to replicate their DMs' decisions accurately.

Result 4. *We find evidence of heterogeneity in the procedures DMs use. As complexity increases, DMs increase their use of some simple procedures, like using only one attribute or writing algorithmic step-by-step messages. We find evidence that DMs who write algorithmic step-by-step messages are more likely to be perfectly replicable. However, these results don't fully drive the increasing gap in replicability as complexity increases.*

ment. For reference, this footnote has the length of the average message in the complex treatment.

[35] Dubra and Ok (2002) propose a similar intuition: "When faced with a more complex problem, what does a decision maker do? An intuitive (and indeed procedural) argument would be that she tries to somehow "break down" the problem into smaller, easier problems that she knows how to solve. If she can really do this (that is, the problem can indeed be decomposed into "easy subproblems"), then she can comfortably make her decision."

[36] This effect holds when additionally coding for words that add second, third, and fourth steps.

[37] Using our coding of messages, separating DMs by those that are perfectly replicable versus those that are not, we find that 36.3% of perfectly-replicable DMs are classified as step-by-step compared to 39.2% of non-perfectly-replicable DMs ($p = 0.315$). Using step-by-step messages as identified by replicators, separating DMs by those that are perfectly replicable versus those that are not, we find that 38.5% of perfectly-replicable DMs are classified as step-by-step compared to 27.7% of non-perfectly-replicable DMs ($p < 0.000$).

III. EXPERIMENT 2: CHARITIES

III.A. Experimental Design

We replicate our main results in a second experiment. This serves as important robustness and extends our results in a few meaningful directions. First, we test our hypothesis in a more naturalistic environment: charity choices. Relative to choice under risk, we don't necessarily have pre-specified hypotheses for procedures that DMs will use in choosing charities, which allows us to test whether procedures emerge naturally in more general choice environments. Second, we test our hypothesis under a new complexity measure—rather than varying complexity by the dimensionality of the object (i.e., the lottery's support size), we instead vary complexity by the cardinality of the *menu*. Specifically, in Experiment 2, DMs choose either from a 2-charity menu or from a 6-charity menu, and we compare the choice processes participants use when choosing from a large or a small menu. The size of the menu is a convenient measure of complexity in that it plausibly varies the cost of decision-making while keeping the objects in the menu the same (up to menu-dependence effects; Kőszegi and Szeidl 2013, Bushong et al. 2021, Somerville 2022). Furthermore, this complexity measure has precedent: The large literature on choice overload shows that decisions from larger menus can come at higher cognitive costs and that DMs respond to this cost by simplifying heuristics, such as choosing the default option (Iyengar and Kamenica 2006).

Other than these differences, the structure of our second experiment follows the general design of our first experiment. Exactly as in our first experiment, our design involves two main types of participants: "decision-makers" and "replicators". We study the choice processes that decision-makers use and use replicators to measure and incentivize the elicitation of the choice process. Screenshots of the instructions for both decision-makers and replicators are included in the Appendix.

Before getting into the details of our experimental design, we highlight an overarching design element that applies to both the decision-maker and replicator studies. By defining menu size as a measure of choice complexity, we assume that choosing from a smaller menu—$\{A,B\}$—is easier than choosing from a larger menu—$\{A,B,C,D,E,F\}$. This need not be true if, for example, the larger menu contains a much better alternative that is not present in the smaller menu. However, as long as $\{C,D,E,F\}$ are irrelevant alternatives, adding them to the menu intuitively makes the decision harder and imposes higher cognitive costs compared to choosing from the smaller menu. The perfect design to identify the effect of complexity, then, would be to conduct a within-subject experiment where we have participants choose their most preferred alternative from a 6-charity menu and then ask them to choose again from a smaller menu that has removed some unchosen charities. Such a within-participant design gives rise to other concerns; namely, it is reasonable to expect that if individuals develop a procedure to choose from the larger menus, then they will maintain the use of that procedure for the smaller menus as well.

Therefore, we use a between-subject design but attempt to maintain the desirable features from the within-subject ideal. We conduct our two treatments sequentially. We first run our complex treatment—where participants see 6-charity menus—and use it to construct the main experiment's 2-charity menus.[38] For each participant and each of the five choices that are to be replicated in the 6-charity menu, we construct a corresponding 2-charity menu as the selected choice and one other charity randomly selected from the remaining five in that menu.[39] We end up with 25 distributions of binary menus, each corresponding to the 25 6-charity menus that participants in the complex treatment see. Our DMs in the simple treatment then get allocated a 2-charity menu drawn, with replacement, from the resulting set of 2-charity menus for each of the 25 unique menus DMs in the complex treatment choose from. Thus, each menu in the simple treatment contains a charity chosen in the corresponding menu in the complex treatment. This improves menu comparability by increasing the likelihood that participants in either treatment pick the same charity. We discuss further benefits of this strategy when addressing the replicators' study.

III.A.1. Decision-Makers Study

To study the use of procedures, we give decision-makers (DMs) a sequence of 25 incentivized choices between charities, presented in random order. Each problem consists of choosing one from a menu of charities. We incentivize these decisions by randomly selecting one chooser and one of their decisions and donating $1,000 to the charity selected by that chooser in that decision. All 25 menus in both treatments are unique, and no charity is repeated across menus.

To study how complexity affects the choice process, we randomly assign participants to one of two treatments that vary the complexity of the choice as defined by the size of the menu DMs choose from. All choosers in the complex treatment face the same 25 menus. Choosers in the simple treatment face 25 menus, each drawn from a different distribution that is the same across participants. In both treatments, the order of the menus is indepen-

[38]These were done at the same time on subsequent weekdays to keep the sample we draw from for each treatment similar.

[39]We draw the randomly selected other charity independently for each subject in the complex treatment, so there is variation in the 2-charity menus even if the chosen charity is the same. Moreover, we only use the five decisions that are to be replicated so that the distribution of menus that replicators see in each treatment is identical.

dently randomized across subjects.

Simple Treatment: All menus have 2 charities.

Complex Treatment: All menus have 6 charities.

We chose to use charities as our choice objects in part because they are defined by many attributes, some of which individuals generally agree on (e.g., efficiency) and others that are more subjective (e.g., the charity's area of work).[40] Specifically, we presented each charity described by nine attributes: area of work, location, administrative expense ratio, program expense ratio, fundraising expense ratio, liabilities to assets ratio, working capital ratio, fundraising efficiency, and program expense growth.[41] We explain the meaning of all attributes to all participants, and they have access to these explanations while making their decisions. We do not show any other information on the charities, including their name. DMs answered three understanding questions about the charities' attributes in general and one question about their incentives in the Message Task.

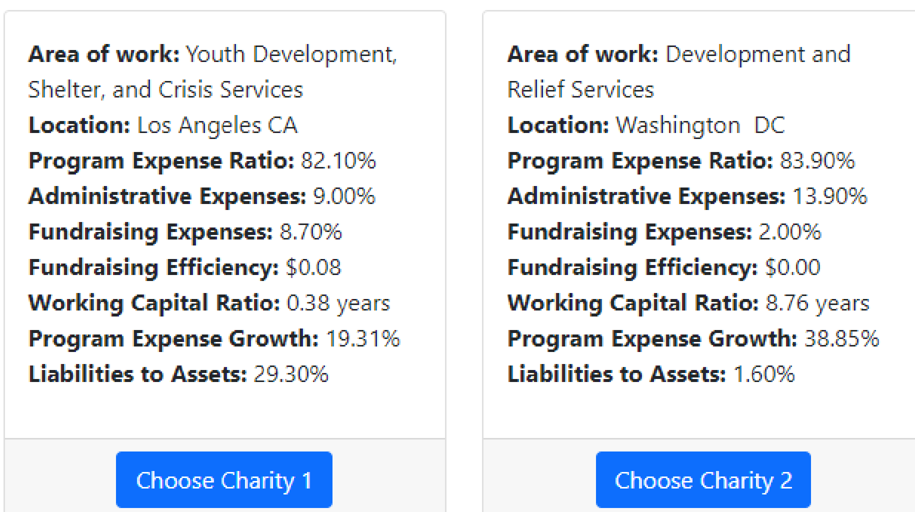


Figure V: Example of charity menu in the Simple treatment

At a random round between rounds 5 and 25, DMs are surprised with the *Message Task*, which works exactly in the same way as it does for Experiment 1.[42] The Message Task elicits their choice process by asking DMs to describe to another participant how they made

[40]Incidentally, this also ensures that no charity in any menu is dominant, which helps ensure that the larger menus are, indeed, more complex, as discussed above.

[41]All charities and their information were taken from `https://www.charitynavigator.org/`.

[42]Part of our pre-registered hypothesis was that procedures would take time to develop, so we did not expect decisions in the complex treatment to be more replicable in early rounds. Thus, in Experiment 1, we included the Message Task only for rounds 10 through 25.

their last five decisions.[43] We incentivize the Message Task by telling DMs that the other participant will see their description of the choice process and try to guess their last five choices and that they earn a $5 bonus payment if the other participant is accurate in a randomly picked guess. They also know the other participant gets the same bonus if the guess is accurate.

Our design again contains features that incentivize DMs to describe their decision-making process rather than individual choices. First, we do not mention charities' names, which plausibly makes the individual charities harder to remember. Second, DMs know, when writing their message, that the other participant will face the five decisions in random order, that we randomize the positioning of the charities on the screen within a decision, and that the charities are randomly renumbered. Finally, the message elicitation comes as a surprise to the DMs, so they have no incentive to attempt to remember their decisions or change their process while choosing.

After submitting their message, we elicit DMs' beliefs about how many decisions, out of five, the replicator will accurately guess. We incentivize this elicitation by randomly picking a DM and giving them an additional $10 if their beliefs are accurate.

III.A.2. Replicators Study

Just as in Experiment 1, the group of participants acting in the role of replicators in our study serves an auxiliary purpose to identify and incentivize the elicitation of the DMs' choice process. We measure the describability of a choice process by the number of choices that someone reading the description accurately guesses. Just as in Experiment 1, each replicator is assigned to either the *NoMessage* or *Message* condition. Again, these conditions only vary in whether the replicator has access to the decision-makers' description of their choice process; in all other ways, the two conditions are identical.

We randomly matched each replicator to three DMs from the same treatment, so each replicator made a total of 15 guesses. Replicators matched to choosers from the Simple Treatment see the exact 2-charity menus that the choosers faced. Replicators matched to choosers from the Complex Treatment *also* see a 2-charity menu: In particular, they see exactly the 2-charity menu that we constructed to create the distribution of menus from which we draw menus for the simple treatment (which contains the DM's chosen charity plus one randomly selected charity from the remaining five). This is an important design feature as it keeps replicators' decisions constant across treatments and ensures a similar

[43] At the time of writing the message, all participants have access to a list of all nine attributes and the definition of these attributes, as well as a list of all possible values that the categorical Area of Work attribute can take to keep information similar across treatments.

"random replication" benchmark.[44] DMs in the complex treatment know that the replicator matched to them will try to guess their choice from a 2-charity menu and that the menu consists of the charity they chose and a randomly picked one from the same menu.[45]

We tell replicators in both treatments about the DM study, and we show replicators the full set of DMs' instructions, including the DMs' instructions about selecting charities in general, the descriptions of all of the charity attributes, and the DMs' instructions during the Message Task for replicators in the Message Condition. For each DM that they are matched to, replicators see the five menus that the DM saw before they faced the Message Task and are asked to select the charity they think the DM chose in that decision. Replicators in the Message condition see, above the two charities, the message that the DM wrote describing their choice process.[46] We incentivize the replication task in both conditions by telling replicators that they may earn a $5 bonus payment if a randomly picked guess is accurate.

III.A.3. Identification

As discussed in Experiment 1, our treatment effect is muted by menus with high agreement. Thus, just as discussed above, our main results will focus on the least "obvious" menus. We identify such menus using choice probabilities in our Simple treatment to create a measure of "menu obviousness," just as we do in experiment 1. According to our measure, the more DMs in the Simple treatment who choose the same charity from the binary menu, the more obvious the menu is.[47] Then, we hypothesize the largest treatment effect for the least obvious menus, where the message is expected to play a more prominent role and where there's a larger scope to identify a treatment effect. Again, our results are qualitatively robust to considering all menus, and we present this in the Appendix and note specific references throughout the text.

Second, we hypothesized that rules take time to develop. To the extent that participants in Complex Treatment only resort to the procedural choice process in later rounds,

[44]Not doing this would severely impact replication rates in the Complex treatment since it's simply easier to guess correctly out of two options than it is to guess correctly out of six options.

[45]This is another reason why DMs don't have a strong incentive to describe the actual charities in the menus they face; A message that says "In the decision with two animal charities, I picked the charity that supported homelessness" is much less helpful to the replicator since they will see only the selected charity and one other.

[46]In both conditions, replicators have access to their instructions and to the descriptions of the charity attributes.

[47]This is an additional feature of designing the lotteries that replicators see when matched to DMs in the complex treatment the way that we do: it allows us to use the choice probabilities from the simple treatment to identify obvious menus in the complex treatment.

we shouldn't expect to observe a treatment effect in early rounds[48]. Moreover, the first rounds will also be confounded if participants in the Simple treatment are better able to recall their last few choices by virtue of having seen fewer alternatives, which would lead to high replicability rates. Our main results still include the first 5 rounds, where participants are plausibly in the process of developing procedures, and we show that excluding the first five rounds, the results are stronger and more statistically significant in Appendix figure XXIII.[49]

III.A.4. Implementation and Recruitment Details

We recruited all participants on Prolific.[50] We recruited 1000 participants from the DMs study across both treatments. Each participant received a \$3 completion payment and took around 16 minutes to complete the study on average. We recruited 708 participants for the replicators study across both conditions. Each replicator also received a \$3 completion payment and took around 11 minutes to complete the study on average.

In all study versions, participants receive ample instructions and are required to correctly answer understanding questions before proceeding to the main parts of our study. Rather than excluding participants, they are given as many times as needed to correctly answer the understanding questions. For full experimental instructions of all study versions that we run, see the supplemental Appendix.

III.B. Experiment 2 Results

Figure VI presents our main result, analogous to Figure II from Experiment 1: We find a message gain of 9.3 pp in the *Simple* treatment, and one of 14.7 pp in the *Complex* treatment.[51] Note that without a message, the accuracy levels are almost the same across treat-

[48]We pre-registered the emergence of a treatment effect in later rounds in AEARCTR-0010977.

[49]As mentioned in Footnote 42, we ran our experiment 1 with lotteries after our experiment 2 with charities. We wanted to test our pre-registered hypothesis that procedures take time to develop, so we allowed the Message Task to show up in rounds 5, 6, 7, 8, and 9 in Experiment 2. Having confirmed our hypothesis, we only allowed the Message Task to show up in rounds 10 and later for experiment 1 in order to increase our power to identify the treatment effect.

[50]We recruited participants in the role of DMs for the complex and simple treatments on March 8^{th} and March 9^{th} of 2023, respectively, and those for the replicators study on March 10^{th}, April 10^{th} and April 14^{th} of 2023. In order to qualify for our study, participants were required to be located in the USA, be fluent in English, and have a minimum of 100 prior submissions on Prolific, with a perfect approval rate. The experiment was implemented using the oTree platform (Chen et al., 2016).

[51]Figure VI shows results for all participants, regardless of the round in which they were surprised with the Message Task. As discussed in our pre-registration (AEARCTR-0010977), we hypothesized that procedures would emerge in later rounds, so we did not expect a treatment difference in earlier rounds, and indeed, there is not one; see Appendix Figure XXIII for details. Because we confirmed our hypothesis in our charity experiment, we did not elicit the message in rounds 5–9 in Experiment 1.

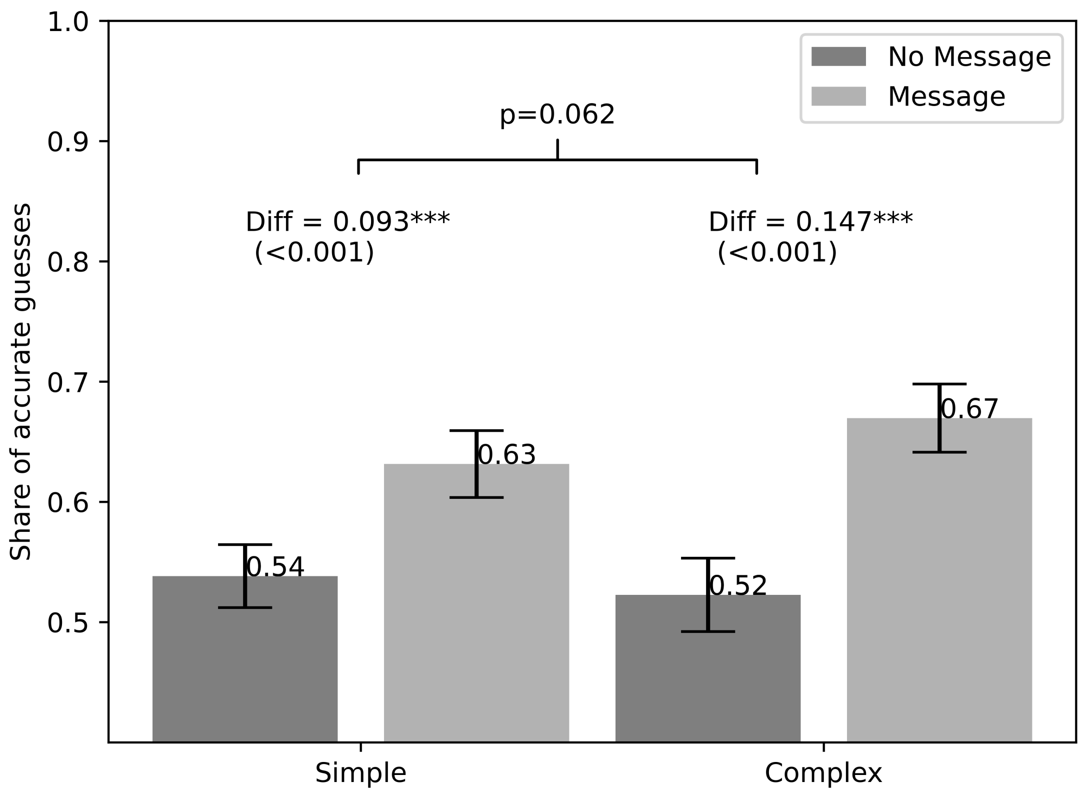


Figure VI: Share of perfectly replicable DMs by treatment and condition.
Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments and conditions. The sample considered is the sample of non-obvious menus, as described in Section III.A; see figure VII in the Appendix for the full sample without this restriction, and figure XXIII for the sample of non-obvious menus split by early and later rounds. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

ments, which is expected since, by design, replicators face basically the same menus across treatments.[52]

Simple Dummy	0.0949
	(0.062)
Complex Dummy	0.0095
	(0.088)
Message Dummy	0.2766
	(0.076)
Message *Complex	0.3168
	(0.085)
Menu Obviousness	0.0074
	(0.001)
Obviousness * Complex	-0.0003
	(0.001)
Obviousness * Message	-0.0026
	(0.001)
Obviousness * Message * Complex	-0.0025
	(0.002)
Surprise Round	-0.0007
	(0.002)
Round * Complex	2.97e-05
	(0.002)
Round * Message	-0.0013
	(0.002)
Round * Message * Complex	0.0026
	(0.003)

Table II: Effect on accuracy of replication
Notes: This report results from an OLS regression. The dependent variable is a binary indicator for the accuracy of the guess. The dummy for simple2 is estimated as the constant in the regression (omitted category). We report standard errors in parentheses; standard errors are clustered at the replicator-level. See appendix table IV for details.

III.B.1. Does procedural decision-making affect the chosen charities?

In our lottery experiment, we considered notions of choice quality and choice consistency. In this experiment, because the choice *objects* are exactly the same across treatments, we can conduct different analyses on decision-making to see whether procedures affect the choices

[52]As a reminder, the only difference between the menus that replicators in each treatment see is that replicators in the simple treatment see random draws with replacement of the menus that replicators in the complex treatment see, where each draw corresponds to one of the 25 unique 6-charity menus the DMs in the complex treatment chose from (this means, for example, that no decision-maker ever sees the same charity twice).

that DMs make. An ideal analysis would have a single individual choose from a 6-charity menu and then from a 2-charity menu that contains their preferred charity from the larger menu (i.e., a test of IIA). This type of within-subject design would raise concerns about choice process contamination, among other issues.

Instead, we estimate a discrete choice model to test whether individuals put different weights on attributes across treatments. We do find differences—specifically, DMs in *Complex* choose charities with lower program expense ratios, lower fundraising expenses, and lower administrative expenses. We do not place any normative weight on this, but just note that the model estimates suggest that individuals are choosing different charities across treatments.[53]

Estimating the model also allows us to test whether the model fits the data better in one treatment versus the other. We might think that a choice model would be more likely to fit procedural decisions: Our choice models are precisely structured algorithms, so decisions made with structured algorithmic approaches might be better approximated by these tools. This exercise also allows us to test the procedural nature of decisions across treatments in a way that bypasses the message description, which provides robustness to our findings. In particular, it shows that our main results using replicability as a measure of describability are not driven by decision-makers being asymmetrically aware of the choice process they are following across treatments.

To test this, we conduct the following analysis, for which we use all of the DMs' data in the simple treatment, and, for the complex treatment, we use the 2-charity menus that replicators see to keep the number of observations used to estimate the models the same across treatments.[54] First, we split each treatment's data in half randomly. We then estimate the model on one-half of the data and predict choices in the other half, given the model estimates. We calculate the root mean square error (RMSE) of these predictions and then repeat this process 100 times. We find that the RMSE is consistently lower in the *Complex* treatment, suggesting that the model does indeed approximate choices better in this treatment (appendix figure XXVIII shows the histogram of the RMSE for both treatments).

III.B.2. Are procedures simplifying the choice process?

As we did in experiment 1, we look for evidence of whether individuals use procedural decision-making to simplify the decision-making process. Again we asked DMs whether

[53]Note that this comparison is meaningful because of the way we design the menus in the simple treatment, by which, in aggregate, the charities chosen in the complex treatment are available in the simple treatment.

[54]This is, each model is estimated using 25000 rows of data, where each row represents an alternative faced by a DM in one of the 25 choice problems.

they used a “shortcut” to make decisions, and we find that 32% of DMs in *Simple* reported using a shortcut while significantly more—41%—in *Complex* did ($p = 0.003$).[55]

As validation of this self-reported measure, we find that DMs who self-report using a shortcut are, indeed, more procedural according to our measure. Specifically, aggregating across treatments, we find that 43% of DMs who reported using a shortcut are perfectly replicable, compared to 35% of those who reported not using a shortcut or using it only sometimes ($p = 0.022$). Thus, we find that DMs who reported using a shortcut are more procedural according to their associated replication rates, and the size of the gap increases after excluding those DMs surprised before round 10.[56]

As a final piece of suggestive evidence, we asked DMs in the follow-up survey how “easy” they found it to choose their preferred charity.[57] We find no significant differences in DMs responses (see Appendix Figure XXVI).[58] Thus, even though DMs in Complex were making decisions 3 times more alternatives, they did not report that the decisions were significantly harder to make. This suggests that the use of simpler choice processes mitigated the cognitive cost of difficult decisions. As validation of this self-reported measure, we find that DMs who self-report finding it “very easy” to pick a charity are, indeed, more procedural according to our measure. Specifically, aggregating across treatments, we find we find that 48% of DMs who reported finding it “very easy” to pick a charity are perfectly replicable, compared to 36% of those who reported finding it “somewhat easy.” ($p = 0.006$), and 34% for those who reported it was somewhat difficult ($p = 0.009$).[59] Thus, we find that DMs who reported finding it easier to pick are more procedural according to their associated replication rates.

III.B.3. What procedures are people using?

We find compelling evidence that individuals are using more procedural approaches in the *Complex* treatment. In choosing charities, we don’t have many pre-specified procedures that individuals are likely using, but we conduct exploratory analysis to find evidence on the types of procedures that individuals use.

Similar to choice under risk, we identify decision-makers whose five choices prior to the

[55]The size of the gap is slightly larger and remains statistically significant after excluding those DMs surprised before round 10 ($p = 0.004$).

[56]Excluding those DMs surprised before round 10 we find that 46% of DMs who reported using a shortcut are perfectly replicable, compared to 36% of those who reported not using a shortcut or using it only sometimes ($p = 0.006$). This difference holds within both treatments.

[57]Specifically, we phrased the question: “Did you find it easy to decide which charity to donate to?”

[58]Excluding rounds before 10 doesn’t change this result.

[59]We also had an option for DMs to report that they found it very difficult. 4 out of the 8 DMs who picked that option are perfectly replicable. Excluding DMs surprised in the first five rounds yields almost identical results.

message task are consistent with maximizing or minimizing a single attribute (e.g., always choosing the charity with the highest program expense ratio) and find no significant differences by treatment or for those perfectly replicable.[60]

Finally, we can again analyze the message data. While we find that messages are of similar length across treatments and unrelated to replicability (see appendix figures XXIV and XXV), we find evidence of step-by-step algorithmic language. Following the same coding above, we find that significantly more messages in Complex are classified as step-by-step (35.6% vs. 41.9%, $p = 0.041$), suggesting that decision processes in complex decisions become more structured and algorithmic.[61] Finally, separating DMs by those that are perfectly replicable versus those that are not, we find that 42.1% of perfectly-replicable DMs are classified as step-by-step compared to only 37.9% of non-perfectly-replicable DMs.

IV. Discussion of Design Considerations

IV..1. Complexity

Our two experiments test our hypothesis using different notions of what makes a decision complex. In general, we think of choice complexity as the cognitive costs associated with choosing one's preferred alternative that stem from the environment of choice.[62] Testing our hypothesis using different complexity measures is crucial for the robustness and external validity of our findings. From a methodological point of view, each measure presents a different set of features. In our risk experiment, we exogenously manipulate complexity by making the objects of choice more complex, in the sense that each consists of more information, which leads to higher information processing costs in characterizing alternatives. Varying the number of outcomes in lotteries allows us to relate to literature in choice under risk that has studied risk preferences in the face of complexity (e.g., Bernheim and Sprenger 2020; Puri 2023) and makes, within treatments, the decision-making exercise directly mimic the replication exercise in our study, without need of adjustments. In our charity experiment, we keep the objects of choice the same (up to menu-dependence effects; Kőszegi and Szeidl 2013, Bushong et al. 2021, Somerville 2022), and exogenously manipulate complexity by increasing the number of alternatives in the menu, which can similarly

[60]Note that this exercise is less informative than it was in the risk experiment given that (i) there are many more distinct attributes describing charities, and (ii) some are subjective, like the charity's location, which makes it impossible for us to test its use from choice data.

[61]This effect holds when additionally coding for words that add second, third, and fourth steps, with treatment levels being 38.8% for Simple and 43.3% for Complex.

[62]Heiner (1983) discussed how the complexity of an environment relates to the agent's competence in deciphering relationships between its behavior and the environment. We abstract from the agent's competence for this study and focus on complexity stemming from the choice environment.

lead to higher information processing costs in characterizing the alternatives. Varying complexity while keeping objects the same is a feature of our empirical identification that allows us to attribute treatment differences in accuracy rates directly to differences in complexity and interpret them as differences in how describable choice processes are. This is by virtue of the fact that using the same objects of choice in all treatments allows us to, by restricting the replication menu in the complex treatment to be the associated 2-charity menu, fix the replication exercise to be the same across treatments. The way in which we construct our 2-charity menus keeps the distribution of replication menus almost identical across treatments so that we can attribute differences in accuracy levels across treatments to the describability of the choice process and not to differences in other elements of the replication exercise. This means that, in principle, our results in the charities experiments can also be interpreted without the need for the No Message condition, given that the exercise in that condition is almost identical across treatments.

IV..2. No Message Condition

We test our hypothesis that the decision-making process becomes more describable as decisions become more complex and, therefore, that the DM's message better describes their choices in complex decisions. We identify choice processes that are easy to describe as those that are easy to replicate based on the description of the choice process. However, comparing replicability rates across treatments at face value introduces a confound—it could be the case that some decisions are simply easy to replicate for reasons that have nothing to do with the choice process. To solve this issue in identification, we isolate the causal effect of the choice process description on replicability rates by comparing replication rates with and without the description in what we call the Message and the No Message conditions.

What is our No Message Condition capturing? Given that this condition operates as a control group, to the extent that replicators can predict the use of procedures even without a message and replicate accurately, we will not be able to capture the use of these procedures in our casual effect. Note that by excluding menus in which a large majority of DMs pick the same alternative (i.e., those that are "obvious"), we reduce the possibility that our No Message condition is capturing predictable procedures, for it is intuitively more likely that procedures are predictable in cases in which more DMs agree on their choice. Even without a message, replicators do significantly better than random guessing in all treatments and experiments, which suggests that, to some extent, they can predict what a DM is doing. However, they do equally well across treatments, which is the crucial element to pay attention to in assessing the role of predictable procedures: To the extent that DMs are using procedures that replicators can predict, and hence are being picked up in our No Message condition, it is not the case that this is happening asymmetrically across treatments.[63] While strictly speaking, our estimates only capture the use of procedures that are, in some sense, not predicted by the replicators without a message, this result serves as a diagnostic that suggests that the hypothesis does not hold true exclusively for non-predictable procedures but plausibly holds true for procedures as a whole.[64]

IV..3. The role of decision-making awareness

We postulate our hypothesis that individuals resort to more procedural decision processes as decisions get more complex regardless of the DM's awareness of this fact. However, our characterization of procedures as describable choice processes makes our identification reliant on the DM's ability to describe the choice process. This means that, in principle, differences in the accuracy of replication across treatments could be driven by differences in DMs' awareness of their choice process.[65] If DMs become more aware of their choice process as complexity increases, they might write descriptions that lead to more accurate replications, even if the underlying choice process is the same.

We have diagnostic tests that suggest DMs' awareness is not driving our estimates of the treatment effect. Firstly, in sections II.C and III.B.1, we show evidence that DMs in different treatments make different choices and that within treatment, those more replicable also make different choices. This strongly suggests that we are not only capturing differences in DM awareness.

Moreover, figures XX and XXVII in the appendix show that, to the extent that DMs are more aware of their choice process in the complex treatment, it does not lead them to believe that their decisions will be more replicable; if anything, in the charities experiment DMs in the complex treatment believe the replicator will guess significantly fewer decisions accurately.[66] If anything, DMs in the charity experiment believe to be less replicable in the complex treatment than in the simple one.

[63] This statement is trivial for the experiments on charities given that, without a message, the replication exercise is almost identical across treatments.

[64] Had we found that without a message accuracy levels where, for example, significantly higher in the simple2 treatment, our results would have to be re-interpreted as pertaining to mostly non-predictable procedures since this would constitute evidence suggestive of DMs switching from more predictable to less predictable procedures as complexity increases. This is not what we find.

[65] Discussions on DMs' awareness of *why* they behaved how they behaved (Nisbett and Wilson, 1977) do not show evidence of DMs becoming more aware of their decision-making process as the complexity of the choice problem increases, which is the asymmetry that would confound our identification.

[66] Note that the link between choice process awareness and replicability beliefs rely on the DM being aware of their awareness.

V. Discussion

Using two experiments, we show that individuals use more procedural—defined to be more "describable"—choice processes as the complexity of the choice environment increases. We provide evidence that suggests that procedural choice processes can lead to different outcomes and of different quality. Furthermore, secondary evidence suggests that these procedural choice processes are less cognitively costly to implement, which makes them especially valuable in complex decisions. Finally, this evidence is robust to different notions of complexity and to different choice environments.

We see many important avenues for future work. First, from a backward-looking perspective, it may be interesting to reassess some of the existing literature in the light of procedural decision-making. For example, which heuristics and well-known biases are "describable?" This question is not only interesting per se but because it can help further discipline our understanding of decision-making shortcuts, what triggers them and their impact on decision quality and consistency, which has been a long-standing issue of debate. For example, in the debate between Gigerenzer and Kahneman and Tversky, in his reply to Kahneman and Tversky (1996) Gigerenzer argues that heuristics "remain vague, undefined, and unspecified with respect both to the antecedent conditions that elicit (or suppress) them and also to the cognitive processes that underlie them" (Gigerenzer, 1996). We see the evidence we put forward in this paper regarding the use of procedural decision-making in the face of complexity as a big contribution in this direction, and further work, both studying procedures and other choice processes that can be captured using the methodology we propose, can further contribute to this cause.

Second, much more work needs to be done to understand procedural decision-making as a category of choice process. When are individuals able to describe how they choose? When are they not? What are the features of the environment—complexity-related and otherwise—that affect this? It would be valuable for future work to consider things like the stakes of the decision, familiarity with the decision environment, etc., as potential features that trigger the use of procedures. For example, rules have been discussed as a resource decision-makers use in the face of weakness of will and in the face of accountability (see Schelling 1985 and Sunstein 2023 for a discussion on procedures in the face of weakness of will, and Slovic 1975, Simonson 1989 and Shafir et al. 1993 for the role of accountability).

Third, it would be interesting to understand how different choice processes affect the decision-makers' "confidence" in their choices. As Simon (1976) discussed procedural rationality, it could be that individuals are most confident in their choices when they are confident in the procedure that they use. However, if individuals believe their procedures to be shortcuts, this would not be the case. Understanding this better is important for understanding whether, e.g., procedural decision-makers are more susceptible to influence by nudges, marketing, or manipulation.

Fourth, it is likely the case that complexity does not always increase procedural decision-making. For example, some forms of complexity could fundamentally overwhelm the decision-maker or could confuse her choice process. As our understanding of complexity increases, so too should our understanding of how the complexity of the decision environment affects the choice processes individuals employ.

Fifth, the ability to elicit incentivized choice process data opens up many doors for future analysis. In environments where the researcher has a candidate choice process in mind, she can code (using natural language techniques, human coders, etc.) for instances of this process in the choice process descriptions. In general environments, using modern language analysis techniques can potentially bring light to features of choice processes that the researcher had not considered, even introducing new considerations that can be modeled theoretically.

References

Abaluck, J. and Gruber, J. (2023). When less is more: Improving choices in health insurance markets. *The Review of Economic Studies*, 90(3):1011–1040.

Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Andreoni, J. and Sprenger, C. (2011). Uncertainty equivalents: Testing the limits of the independence axiom. Technical report, National Bureau of Economic Research.

Andreoni, J. and Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, 102(7):3357–3376.

Angeletos, G.-M., Laibson, D., Repetto, A., Tobacman, J., and Weinberg, S. (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *Journal of Economic perspectives*, 15(3):47–68.

Augenblick, N. and Nicholson, S. (2016). Ballot position, choice fatigue, and voter behaviour. *The Review of Economic Studies*, 83(2):460–480.

Aumann, R. J. (1997). Rationality and bounded rationality. *Games and Economic behavior*, 21(1-2):2–14.

Aumann, R. J. (2019). A synthesis of behavioural and mainstream economics. *Nature human behaviour*, 3(7):666–670.

Banovetz, J. and Oprea, R. (2023). Complexity and procedural choice. *American Economic Journal: Microeconomics*, 15(2):384–413.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, 91(1):79–98.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142.

Berger, C. C., Dennehy, T. C., Bargh, J. A., and Morsella, E. (2016). Nisbett and wilson (1977) revisited: The little that we can know and can tell. *Social Cognition*, 34(3):167–195.

Bernheim, B. D. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *American economic review*, 94(5):1558–1590.

Bernheim, B. D., Royer, R., and Sprenger, C. (2022). Robustness of rank independence in risky choice. In *AEA Papers and Proceedings*, volume 112, pages 415–420. American Economic Association.

Bernheim, B. D. and Sprenger, C. (2020). On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting. *Econometrica*, 88(4):1363–1409.

Bonder, T., Erev, I., Ludvig, E. A., and Roth, Y. (2023). The common origin of both oversimplified and overly complex decision rules. *Journal of Behavioral Decision Making*, page e2321.

Brandts, J. and Charness, G. (2003). Truth or consequences: An experiment. *Management Science*, 49(1):116–130.

Bushong, B., Rabin, M., and Schwartzstein, J. (2021). A model of relative thinking. *The Review of Economic Studies*, 88(1):162–191.

Camerer, C. F. and Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2):167–196.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from indonesia. *Economic Inquiry*, 37(1):47–59.

Caplin, A., Csaba, D., Leahy, J., and Nov, O. (2020). Rational inattention, competitive supply, and psychometrics. *The Quarterly Journal of Economics*, 135(3):1681–1724.

Caplin, A. and Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203.

Caplin, A., Dean, M., and Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094.

Caplin, A., Dean, M., and Leahy, J. (2022). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715.

Caplin, A., Dean, M., and Martin, D. (2011). Search and satisficing. *American Economic Review*, 101(7):2899–2922.

Cason, T. N. and Mui, V.-L. (2019). Individual versus group choices of repeated game strategies: A strategy method approach. *Games and Economic Behavior*, 114:128–145.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Cherepanov, V., Feddersen, T., and Sandroni, A. (2013). Rationalization. *Theoretical Economics*, 8(3):775–800.

Chernev, A., Böckenholt, U., and Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2):333–358.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2006). Substantive and procedural rationality in decisions under uncertainty. *Unpublished*.

Cooper, D. J. and Dutcher, E. G. (2011). The dynamics of responder behavior in ultimatum games: a meta-study. *Experimental Economics*, 14:519–546.

Cooper, D. J. and Stockman, C. K. (2002). Fairness and learning: an experimental examination. *Games and Economic Behavior*, 41(1):26–45.

Dal Bó, P. and Fréchette, G. R. (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*, 109(11):3929–3952.

Dean, M. and Neligh, N. N. (2017). Experimental tests of rational inattention. *Unpublished*.

Della Vigna, S. and Malmendier, U. (2006). Paying not to go to the gym. *american economic Review*, 96(3):694–719.

Dubra, J. and Ok, E. A. (2002). A model of procedural decision making in the presence of risk. *International Economic Review*, 43(4):1053–1080.

Enke, B., Graeber, T., and Oprea, R. (2023). Complexity and time. Technical report, National Bureau of Economic Research.

Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1:615–691.

Fisman, R., Kariv, S., and Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, 97(5):1858–1876.

Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3):347–369.

Fudenberg, D. and Puri, I. (2022). Simplicity and probability weighting in choice under risk. In *AEA Papers and Proceedings*, volume 112, pages 421–25.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4):1043–1068.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to kahneman and tversky. *Psychological Review*.

Gigerenzer, G. and Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62:451–482.

Gigerenzer, G. and Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4):650.

Gigerenzer, G. and Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gigerenzer, G. and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2009). *The complexity of the consumer problem and mental accounting*. Pinhas Sapir Center for Development, Tel Aviv University.

Gilboa, I. and Schmeidler, D. (1995). Case-based decision theory. *The quarterly Journal of economics*, 110(3):605–639.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The quarterly journal of economics*, 115(3):811–846.

Gneezy, U., List, J. A., and Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics*, 121(4):1283–1309.

Halevy, Y. and Feltkamp, V. (2005). A bayesian approach to uncertainty aversion. *The Review of Economic Studies*, 72(2):449–466.

Halevy, Y. and Mayraz, G. (2022). Identifying Rule-Based Rationality. *The Review of Economics and Statistics*, pages 1–44.

Heiner, R. A. (1983). The origin of predictable behavior. *The American economic review*, 73(4):560–595.

Holt, C. A. (1986). Preference reversals and the independence axiom. *The American Economic Review*, 76(3):508–515.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9(1):90–98.

Iyengar, S. S. and Kamenica, E. (2006). Choice overload and simplicity seeking. *University of Chicago Graduate School of Business Working Paper*, 87:1–27.

Johnson, E. J., Schulte-Mecklenbeck, M., and Willemsen, M. C. (2008). Process models deserve process data: Comment on brandstätter, gigerenzer, and hertwig (2006). *Psychological Review*.

Kagel, J. H., Kim, C., and Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior*, 13(1):100–110.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

Kahneman, D. and Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*.

Kahneman, W. (1979). D.# and tvereky, a.. prospect theory# an analysis of decision under risk^. *Econometrica*, 47(2).

Kőszegi, B. and Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly journal of economics*, 128(1):53–104.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory*, 170:70–85.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97(387):666–684.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597.

Luce, R. D. (1978). Lexicographic tradeoff structures. *Theory and Decision*, 9(2):187.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2023). Rational inattention: A review. *Journal of Economic Literature*, 61(1):226–273.

Mandler, M., Manzini, P., and Mariotti, M. (2012). A million answers to twenty questions: Choosing by checklist. *Journal of Economic Theory*, 147(1):71–92.

Manzini, P. and Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review*, 97(5):1824–1839.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5):2183–2205.

Mononen, L. (2022). On preference for simplicity and probability weighting. Technical report, Working paper.

Nisbett, R. E. and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, 84(3):231.

Oprea, R. (2020). What makes a rule complex? *American economic review*, 110(12):3913–51.

Oprea, R. (2022). Simplicity equivalents. *Working Paper*.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of experimental psychology: Learning, Memory, and Cognition*, 14(3):534.

Puri, I. (2018). Preference for simplicity. *Available at SSRN 3253494*.

Puri, I. (2023). Simplicity and risk. *Unpublished*.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101(2):900–926.

Romero, J. and Rosokha, Y. (2018). Constructing strategies in the indefinitely repeated prisoner's dilemma game. *European Economic Review*, 104:185–219.

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the allais paradox?). *Journal of economic theory*, 46(1):145–153.

Rubinstein, A. (1998). *Modeling bounded rationality*. MIT press.

Rubinstein, A. and Salant, Y. (2006). A model of choice from lists. *Theoretical Economics*, 1(1):3–17.

Salant, Y. (2011). Procedural analysis of choice rules with applications to bounded rationality. *American Economic Review*, 101(2):724–748.

Salant, Y. and Rubinstein, A. (2008). (a, f): choice with frames. *The Review of Economic Studies*, 75(4):1287–1296.

Salant, Y. and Spenkuch, J. L. (2022). Complexity and choice. Technical report, National Bureau of Economic Research.

Scheibehenne, B., Greifeneder, R., and Todd, P. M. (2010). Can there ever be too many options? a meta-analytic review of choice overload. *Journal of consumer research*, 37(3):409–425.

Schelling, T. C. (1985). Enforcing rules on oneself. *The Journal of Law, Economics, and Organization*, 1(2):357–374.

Schulte-Mecklenbeck, M., Johnson, J. G., Böckenholt, U., Goldstein, D. G., Russo, J. E., Sullivan, N. J., and Willemsen, M. C. (2017). Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science*, 26(5):442–450.

Shafir, E., Simonson, I., and Tversky, A. (1993). Reason-based choice. *Cognition*, 49(1-2):11–36.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, pages 99–118.

Simon, H. A. (1976). From substantive to procedural rationality. *25 years of economic theory: Retrospect and prospect*, pages 65–86.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research*, 16(2):158–174.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Slonim, R. and Roth, A. E. (1998). Learning in high stakes ultimatum games: An experiment in the slovak republic. *Econometrica*, pages 569–596.

Slovic, P. (1975). Choice between equally valued alternatives. *Journal of Experimental Psychology: Human perception and performance*, 1(3):280.

Smith, E. R. and Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to nisbett and wilson. *Unpublished*.

Somerville, J. (2022). Range-dependent attribute weighting in consumer choice: An experimental test. *Econometrica*, 90(2):799–830.

Sunstein, C. R. (2023). *Decisions about Decisions: Practical Reason in Ordinary Life*. Cambridge University Press.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics letters*, 8(3):201–207.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157):1124–1131.

Tversky, A. and Kahneman, D. (1986). The framing of decisions and the evaluation of prospects. In *Studies in Logic and the Foundations of Mathematics*, volume 114, pages 503–520. Elsevier.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5:297–323.

Von Neumann, J. and Morgenstern, O. (1944). Theory of games and economic behavior princeton. *Princeton University Press*, 1947:1953.

Wakker, P. P. (2022). A criticism of bernheim & sprenger's (2020) tests of rank dependence. *Available at SSRN 4188758*.

Weber, E. and Kirsner, B. (1997). Reasons for rank-dependent utility evaluation. *Journal of Risk and Uncertainty*, 14:41–61.

White, P. (1980). Limitations on verbal reports of internal events: A refutation of nisbett and wilson and of bem. *Unpublished*.

Wu, G. (1994). An empirical test of ordinal independence. *Journal of Risk and Uncertainty*, 9:39–60.

A. Additional Figures and Tables

Dep. Variable:	Accuracy	**R-squared:**	0.043
Model:	OLS	**Adj. R-squared:**	0.042
Method:	Least Squares	**F-statistic:**	56.78
Prob (F-statistic):	2.31e-109	**Log-Likelihood:**	-9522.0
No. Observations:	14445	**AIC:**	1.907e+04
Df Residuals:	14431	**BIC:**	1.918e+04
Df Model:	13	**Covariance Type:**	cluster

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
Simple2 Dummy	0.0178	0.039	0.452	0.651	-0.059	0.095
Simple3 Dummy	-0.0406	0.016	-2.601	0.009	-0.071	-0.010
Complex Dummy	-0.0189	0.015	-1.265	0.206	-0.048	0.010
Message Dummy	0.0772	0.076	1.014	0.310	-0.072	0.226
Message * Simple3	0.0379	0.098	0.389	0.698	-0.153	0.229
Message * Complex	0.3168	0.102	3.117	0.002	0.118	0.516
Obviousness	0.9320	0.044	21.064	0.000	0.845	1.019
Obviousness * Simple3	-0.0274	0.053	-0.515	0.606	-0.131	0.077
Obviousness * Complex	-0.2086	0.054	-3.859	0.000	-0.314	-0.103
Obviousness * Message	-0.1329	0.079	-1.677	0.093	-0.288	0.022
Obviousness * Message * Complex	-0.2086	0.054	-3.859	0.000	-0.314	-0.103
Obviousness * Message * Simple3	-0.0274	0.053	-0.515	0.606	-0.131	0.077
Surprise Round	-0.0021	0.001	-1.646	0.100	-0.005	0.000
Round * Simple3	0.0008	0.002	0.492	0.623	-0.002	0.004
Round * Complex	0.0011	0.002	0.645	0.519	-0.002	0.004
Round * Message	0.0016	0.003	0.597	0.550	-0.004	0.007
Round * Message * Simple3	0.0008	0.002	0.492	0.623	-0.002	0.004
Round * Message * Complex	0.0011	0.002	0.645	0.519	-0.002	0.004

Table III: OLS Regression Results

Notes: The sample considers all guesses all replicators make. Standard Errors are robust to cluster correlation at the replicator level.

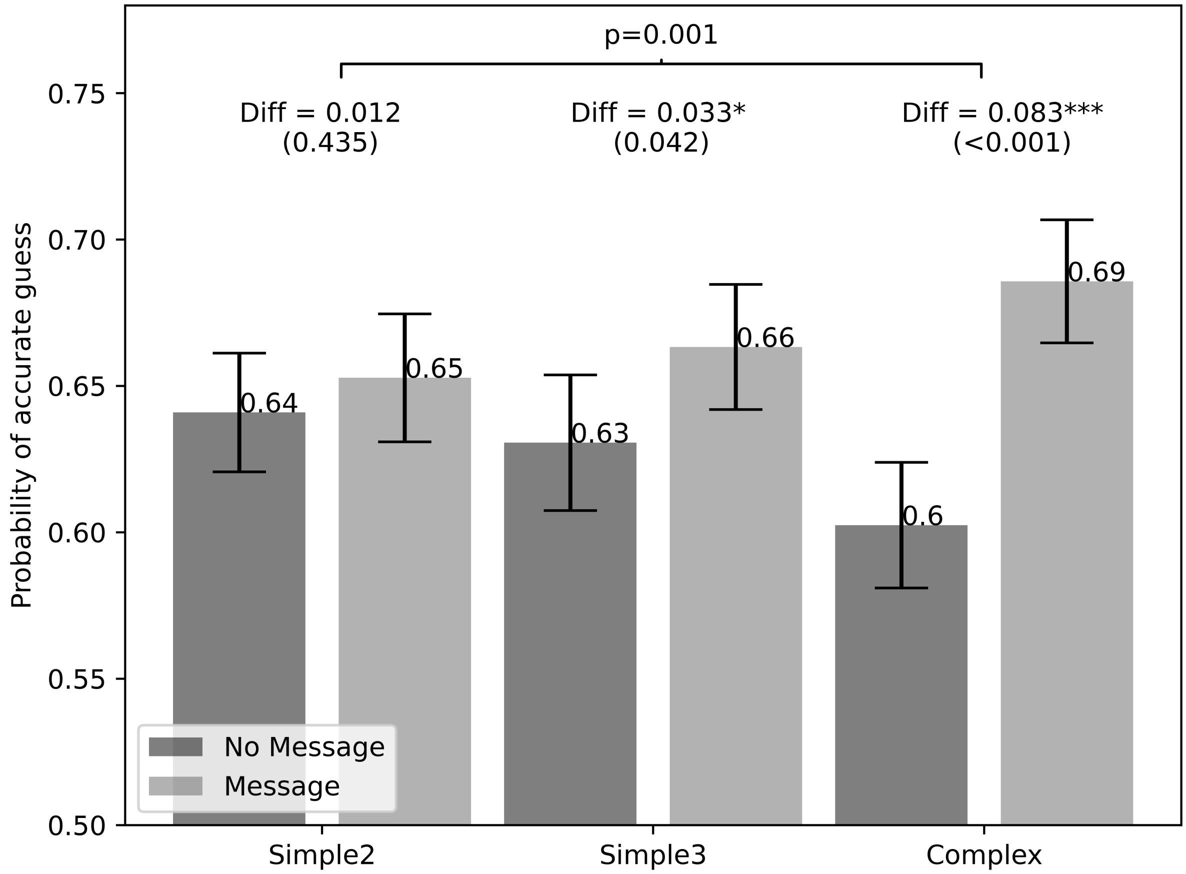


Figure VII: Replication accuracy by treatment and condition.
Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments. The sample considered is the full sample without any restriction; see Figure II in the main body of the paper for the sample with only non-obvious menus, as described in Section II.A.

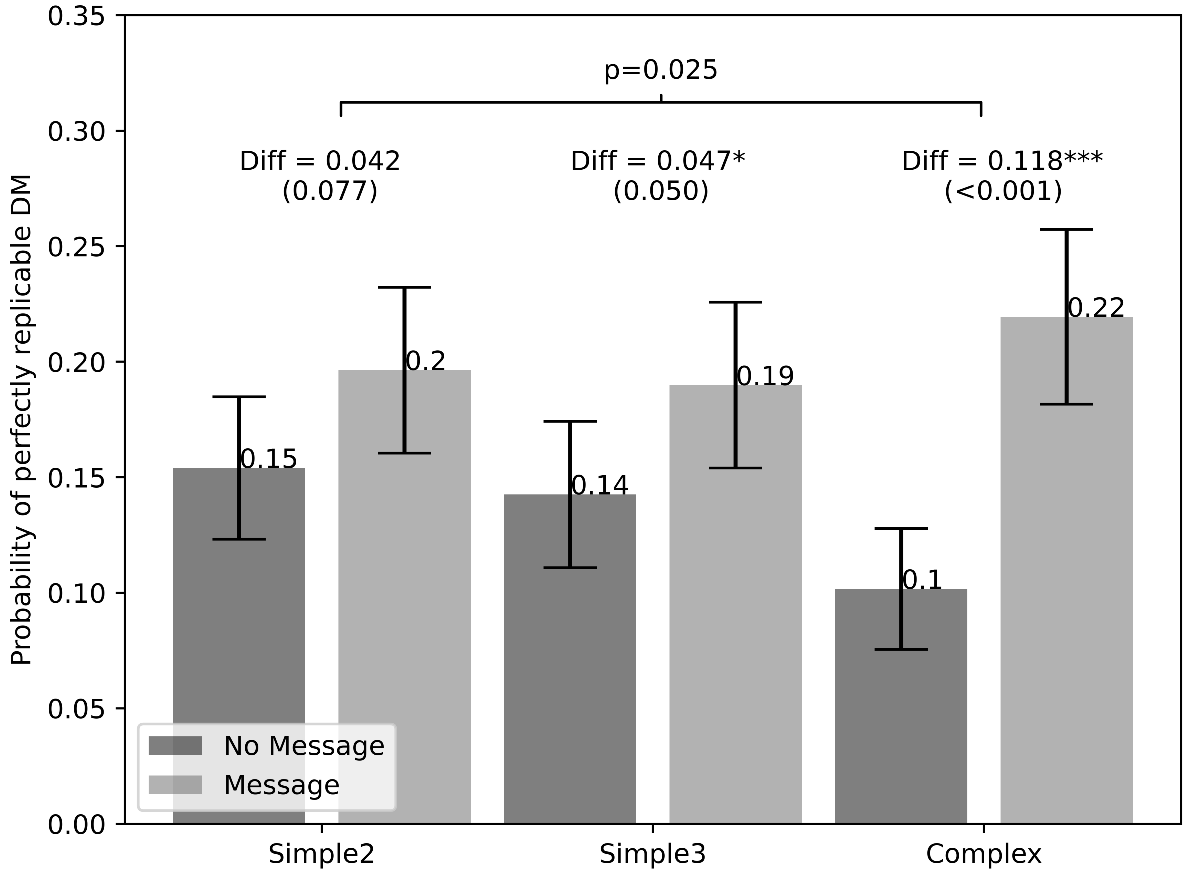


Figure VIII: Share of perfectly replicable DMs by treatment and condition.
Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments and conditions. The sample considered is the full sample; see Figure III in the main body of the paper for the sample restricted to non-obvious menus, as described in Section II.A. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

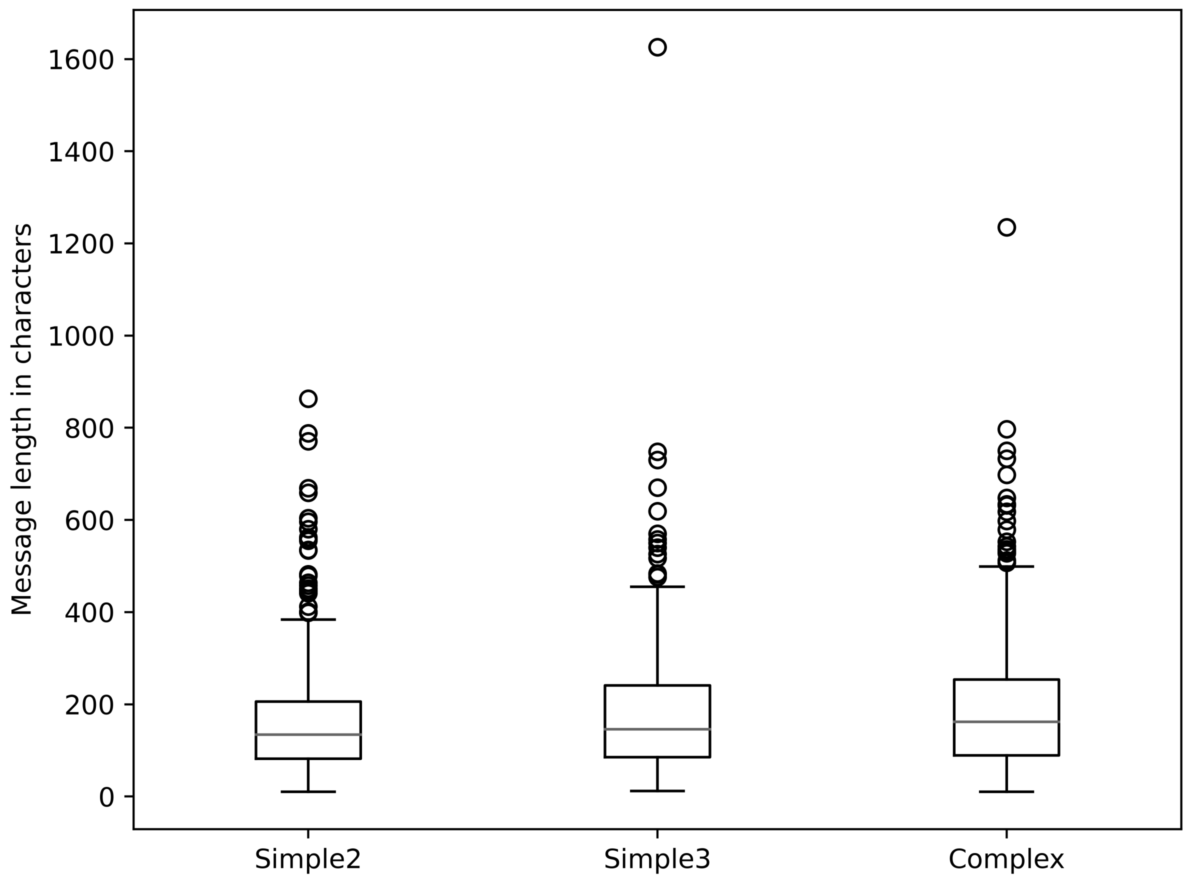


Figure IX: Boxplot of message length by treatment

Note: Median lengths are 134, 146 and 162 for the simple2, simple3 and complex treatments, respectively.

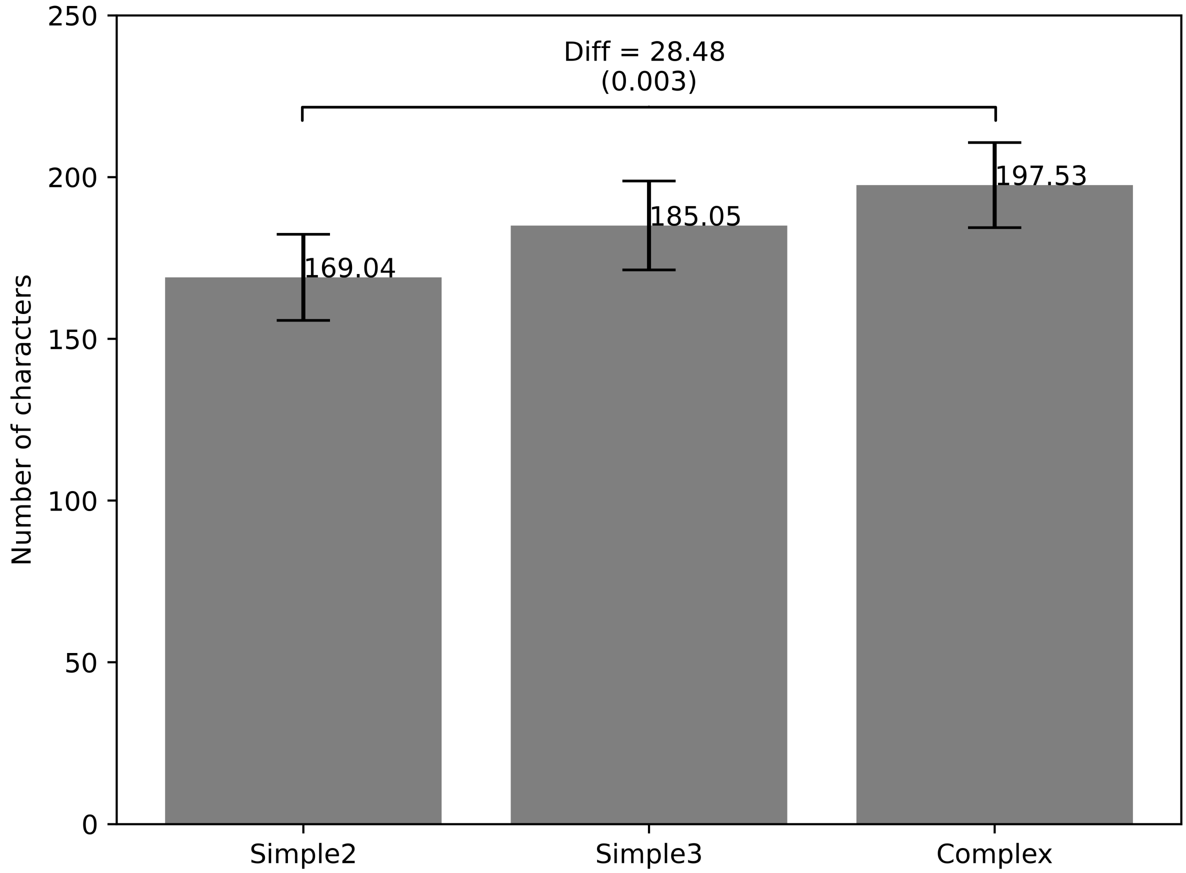


Figure X: Message length by treatment

Note: Bars show the average length of messages in number of characters for each treatment. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions. In parentheses, we show p-values from a t-test at the 95% level.

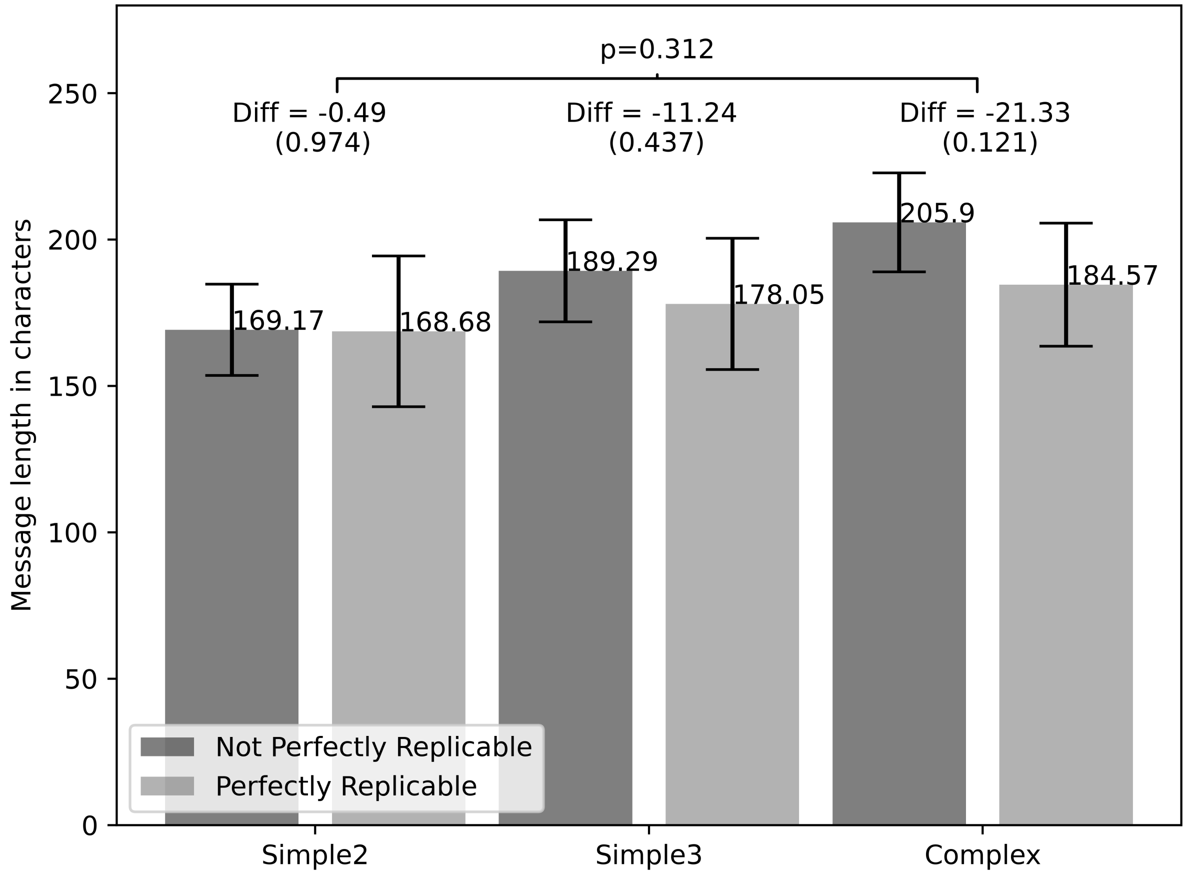


Figure XI: Message length in characters, for each treatment, split by perfectly and non-perfectly replicable DMs.

Note: Bars show the average length of messages in number of characters for each treatment, split by perfectly and non-perfectly replicable DMs. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions. In parentheses, we show p-values from a t-test at the 95% level.

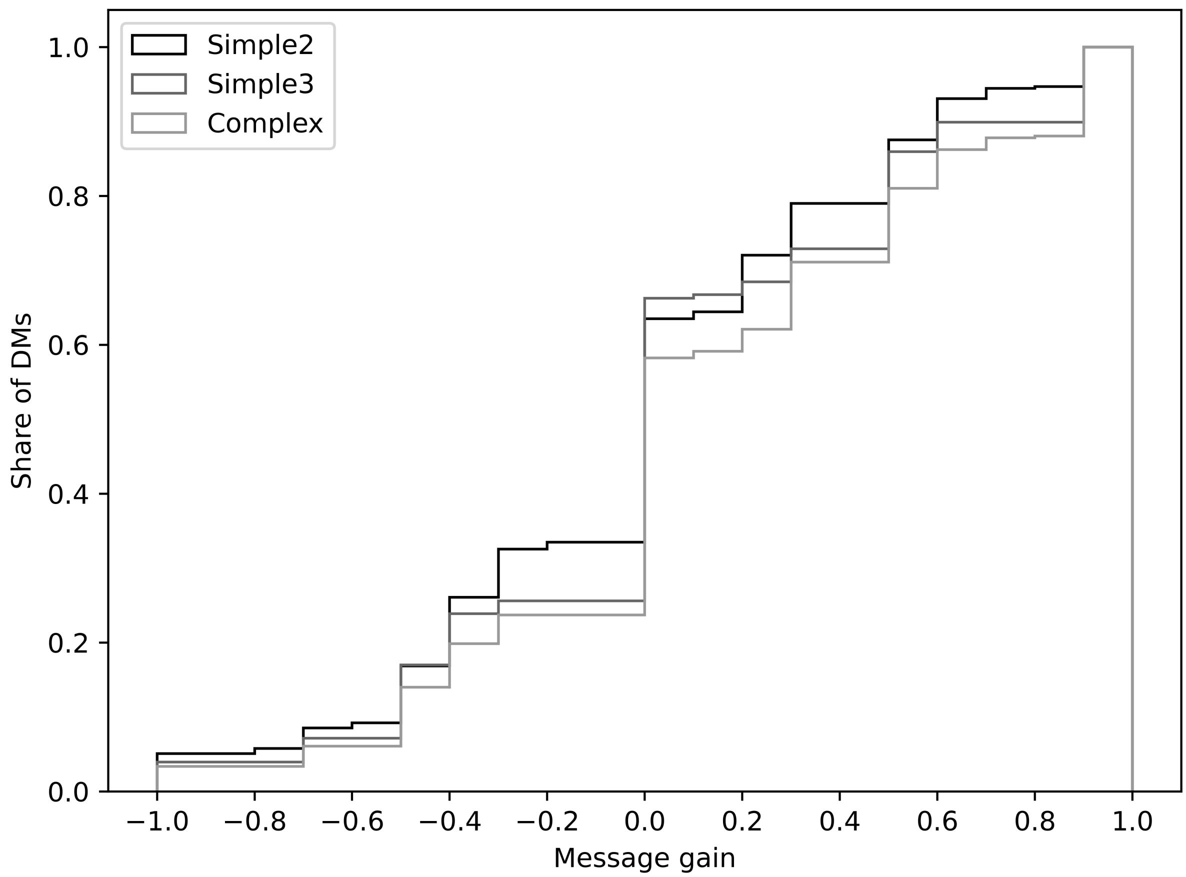


Figure XII: CDF of message gain by treatment.

Note: Message gain is constructed by taking, for each DM, the difference between accuracy levels in the no message condition and in the message condition. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions.

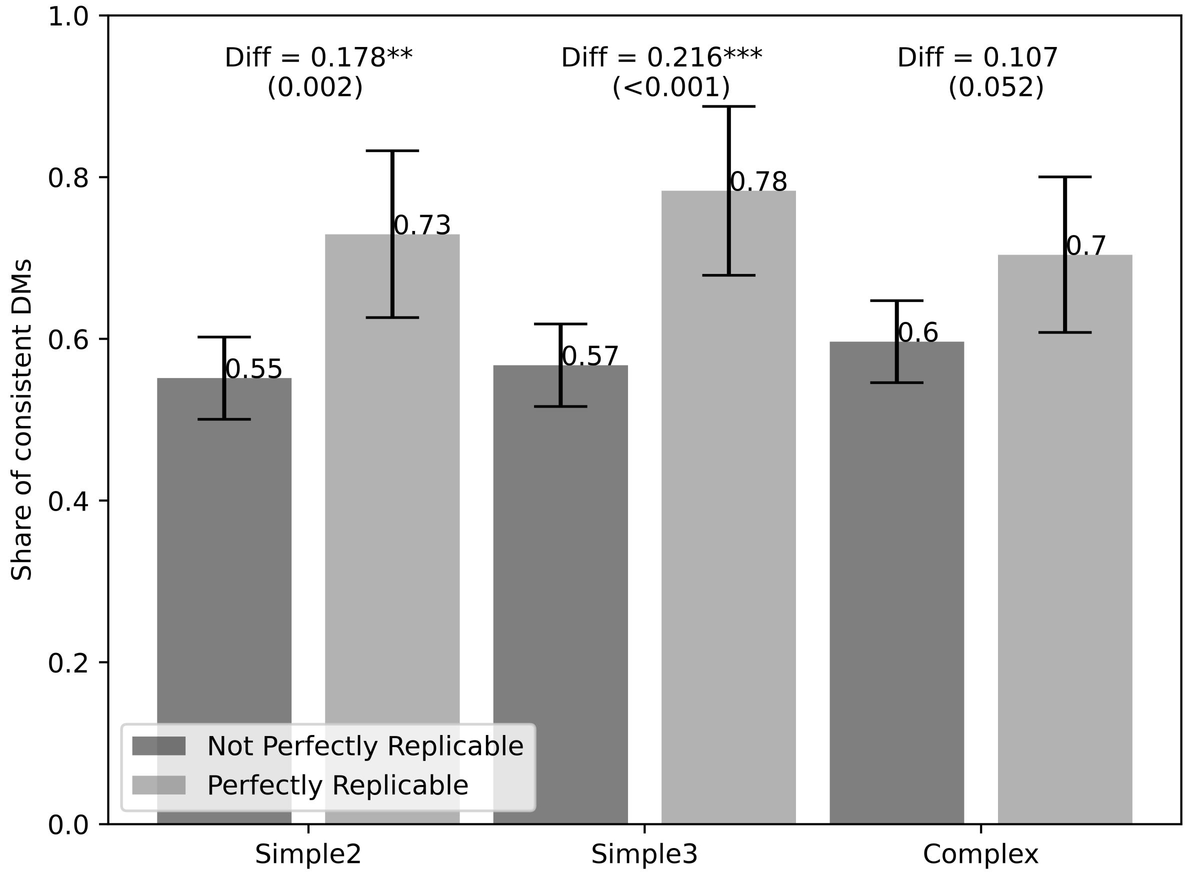


Figure XIII: Share of DMs who make the same choice in all repeated menus by perfectly replicable status and treatment.

Note: Bars show the share of DMs who make the same choice in all repeated menus for each treatment, split by whether or not they are perfectly replicable with a message. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions. In parentheses, we show p-values from a t-test at the 95% level.

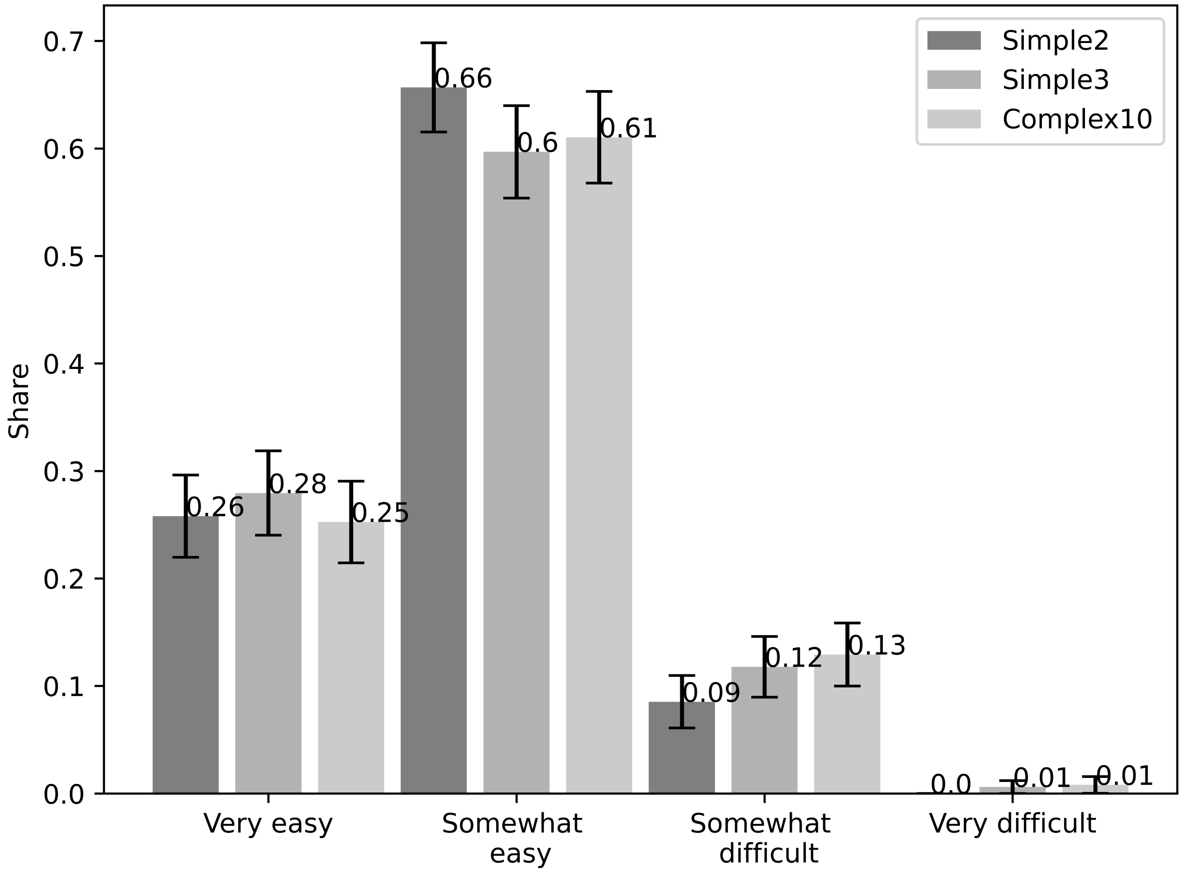


Figure XIV: DMs' responses to the question "How easy was it for you to decide which lottery to choose?"

Note: Bars show the share of DMs who gave each of the responses in the x-axis to the question "How easy was it for you to decide which lottery to choose?" in each treatment. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions. In parentheses, we show p-values from a t-test at the 95% level.

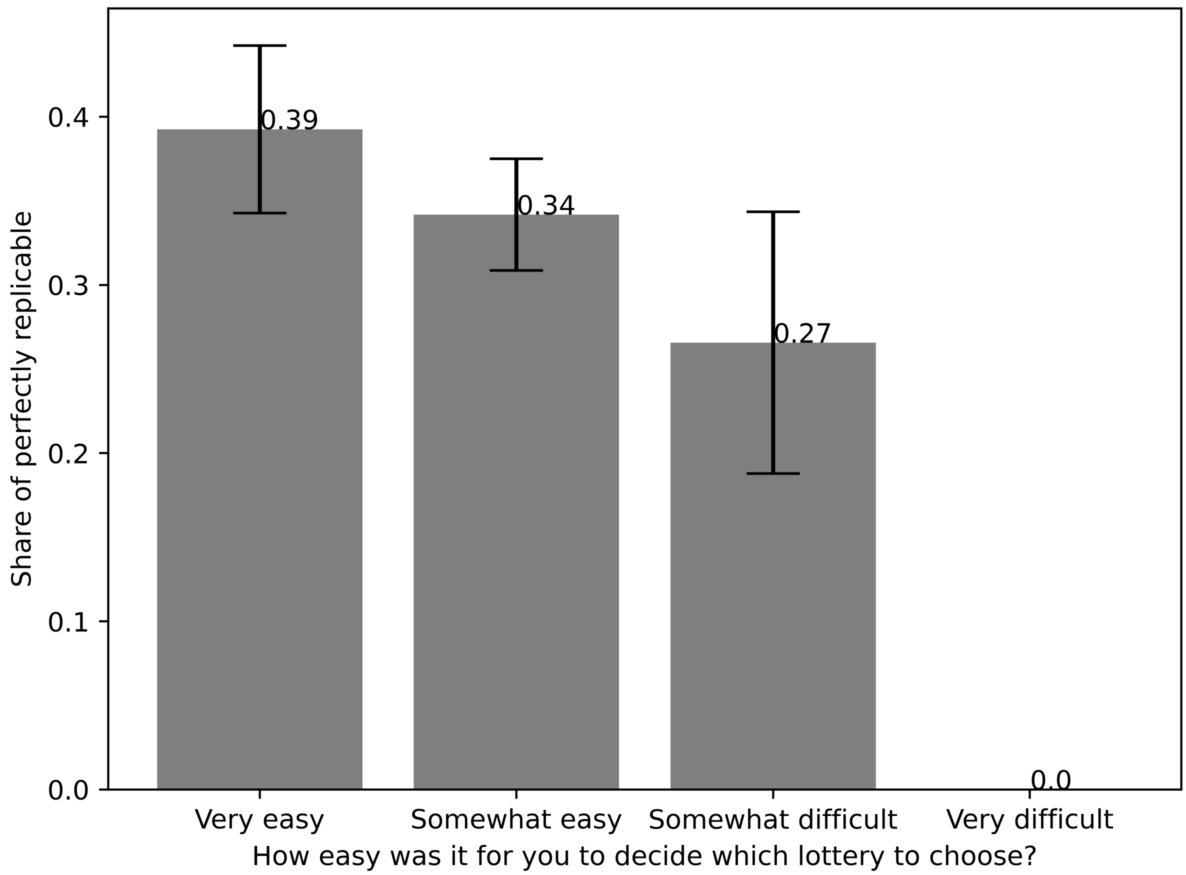


Figure XV: DMs' responses to the question "How easy was it for you to decide which lottery to choose?"

Note: Bars show the share of DMs who are perfectly replicable by each of the responses in the x-axis to the question "How easy was it for you to decide which lottery to choose?" The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions. In parentheses, we show p-values from a t-test at the 95% level.

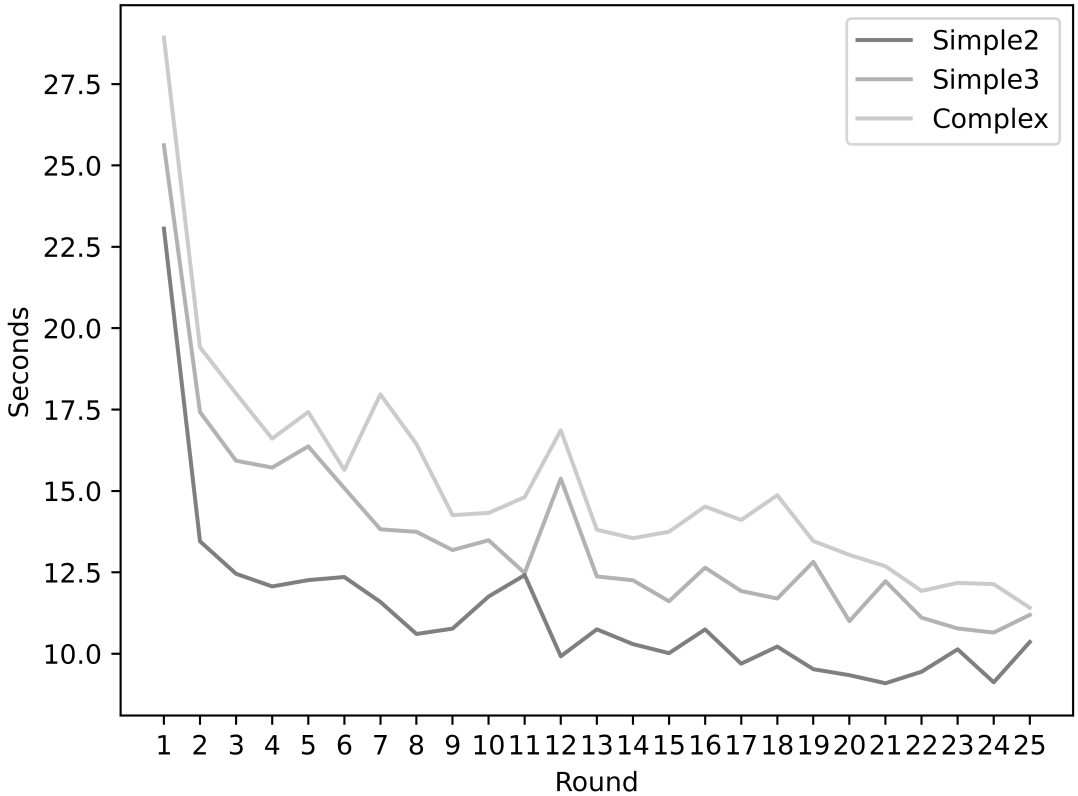


Figure XVI: Time DMs spend choosing

Note: Lines show the time DMs take to pick lotteries, in each round, in seconds, by treatment. This sample considers all DMs.

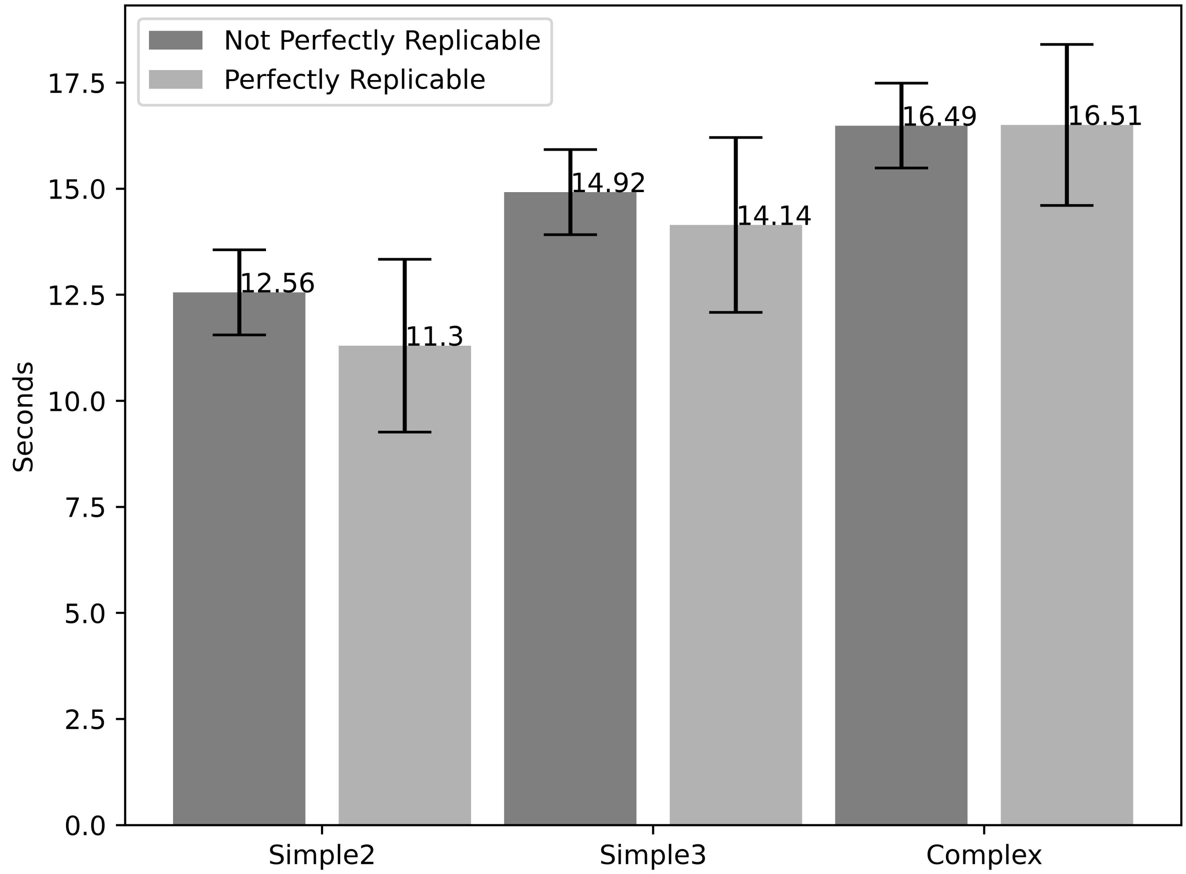


Figure XVII: Time DMs spend choosing by perfectly replicable status

Note: Bars show the time DMs take to guess in seconds by treatment and perfectly replicable status. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

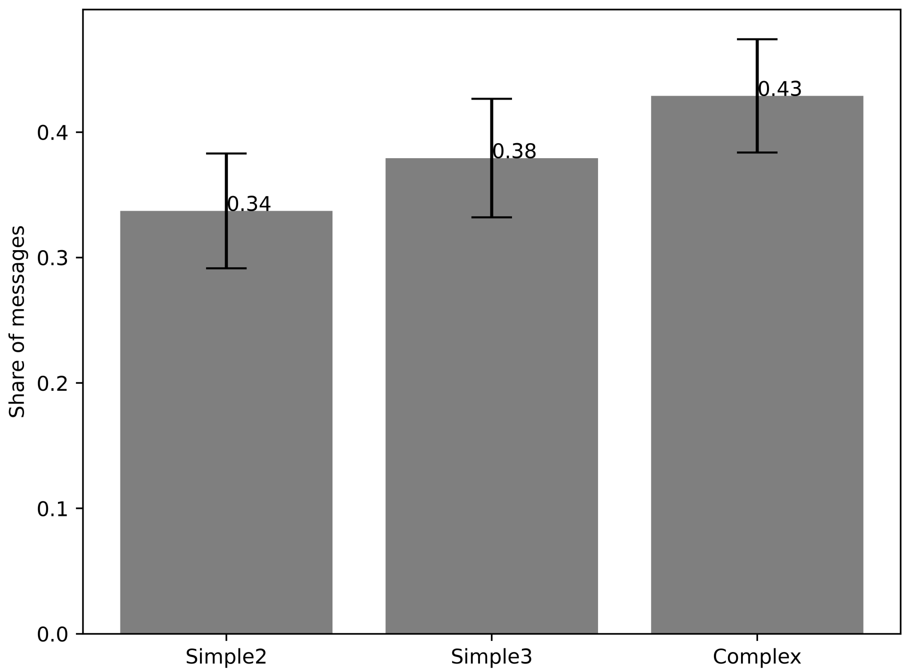


Figure XVIII: Use of algorithmic language in messages.
Note: Bars show the share of DMs' that use algorithmic language in their messages by treatment. The exact words coded for are: 'first', '1)', '1. ', ' one ', ' one.', ' One', 'First ', ' then', and 'Then '. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

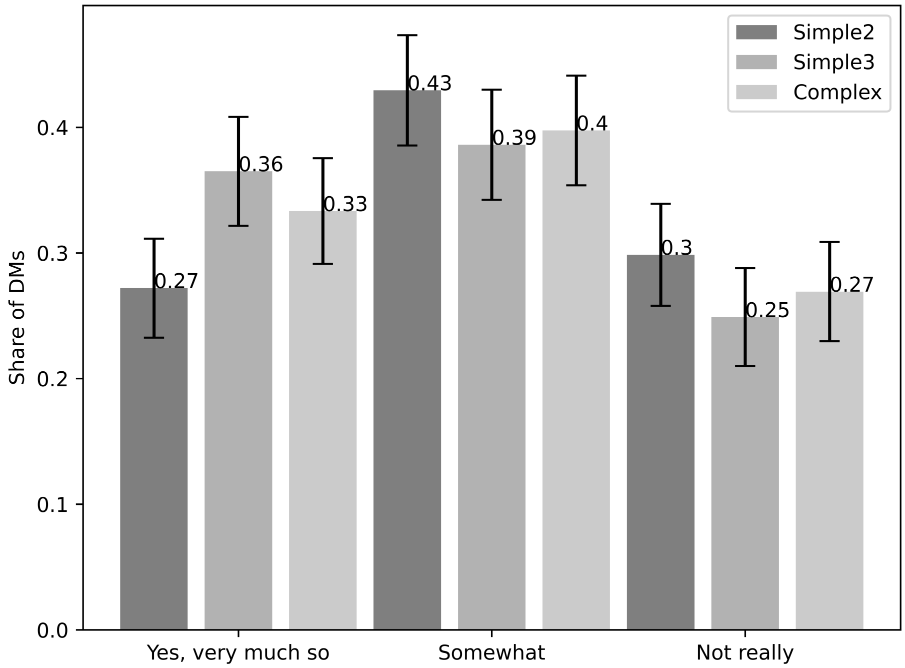


Figure XIX: Replicators' responses to "Did the message feel like a step-by-step (or single-step) process?"

Note: Bars show the share of DMs classified by replicators in the message condition into each of the responses in the x-axis. We ask replicators this question once per DM they see: "Did the message feel like a step-by-step (or single-step) process?" Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

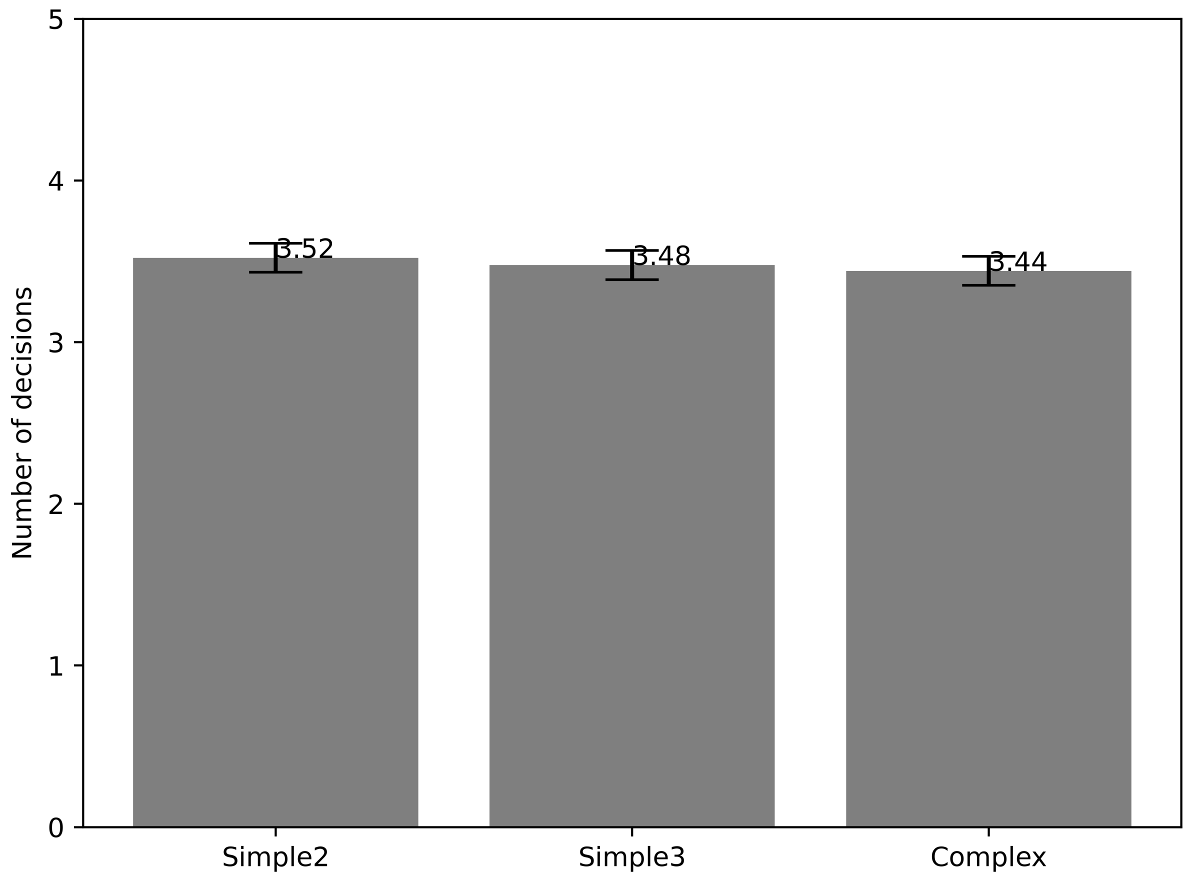


Figure XX: DMs responses to: “Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?”

Note: Bars show the share of DMs that give each of the responses in the x-axis. We ask DMs this question once right after writing their message: "Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?" Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

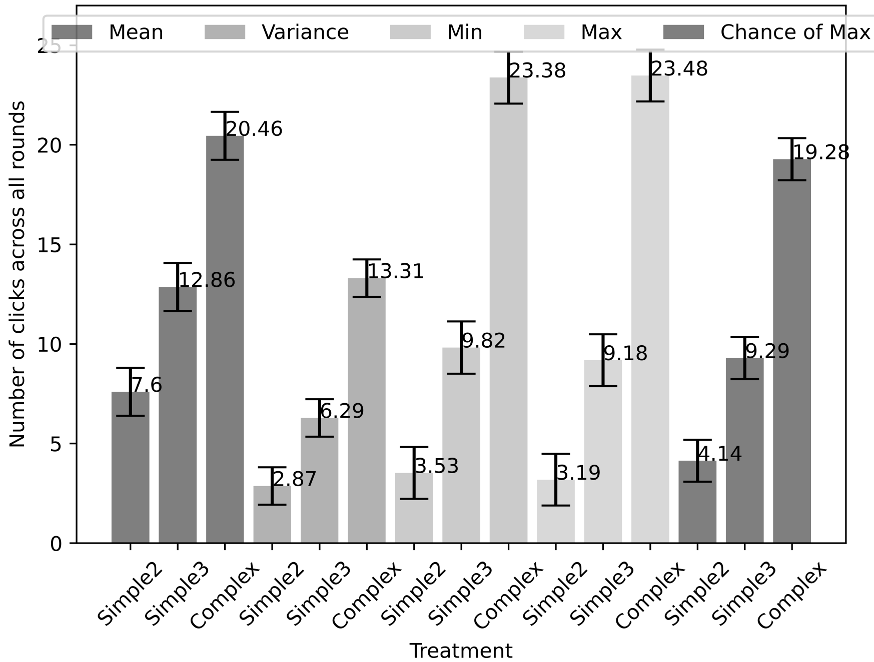


Figure XXI: Average Button Usage by DMs

Note: Bars show the average number of times DMs click on each button throughout their 25 decisions, by treatment, for each of our five buttons separately. The sample considers all DMs. Vertical lines reflect 95% confidence intervals.

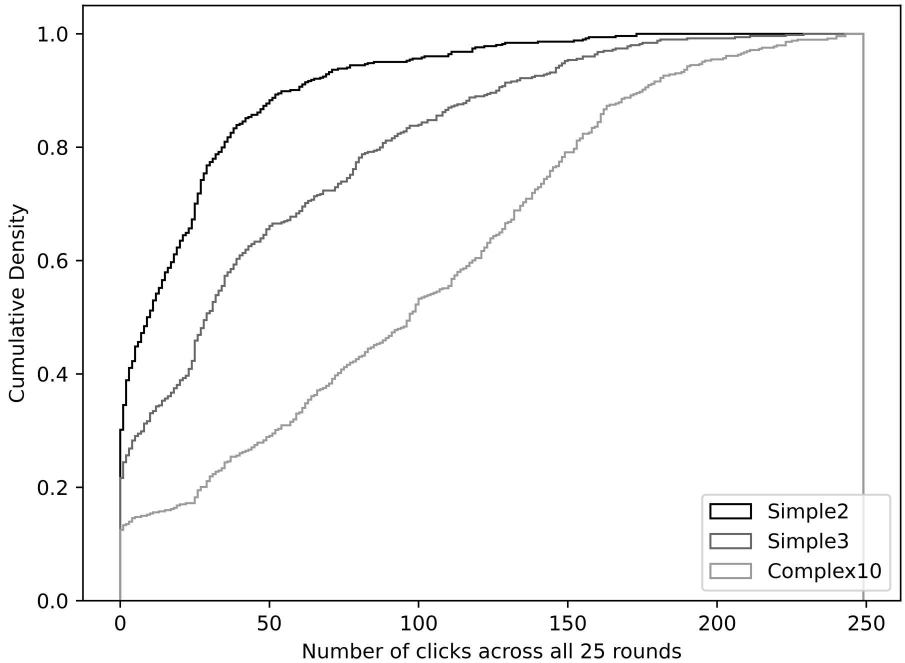


Figure XXII: CDF of Button Usage by DMs
Note: Lines show the CDF for the number of times DMs click on any button throughout their 25 decisions, by treatment. The sample considers all DMs.

A.B. Charities experiment

Dep. Variable:	Accuracy	**R-squared:**	0.029
Model:	OLS	**Adj. R-squared:**	0.028
Method:	Least Squares	**F-statistic:**	25.74
Prob (F-statistic):	4.17e-45	**Log-Likelihood:**	-7015.5
No. Observations:	10620	**AIC:**	1.406e+04
Df Residuals:	10608	**BIC:**	1.414e+04
Df Model:	11	**Covariance Type:**	cluster

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
Simple Treatment	0.0949	0.062	1.536	0.125	-0.026	0.216
Complex Treatment	0.0095	0.088	0.108	0.914	-0.163	0.182
Message Dummy	0.2766	0.085	3.270	0.001	0.111	0.442
Message * Complex	0.1557	0.118	1.315	0.188	-0.076	0.388
Obviousness	0.0074	0.001	9.299	0.000	0.006	0.009
Obviousness * Complex	-0.0003	0.001	-0.300	0.764	-0.003	0.002
Obviousness * Message	-0.0026	0.001	-2.413	0.016	-0.005	-0.000
Obviousness * Complex * Message	-0.0025	0.002	-1.646	0.100	-0.005	0.000
Surprise Round	-0.0007	0.002	-0.425	0.671	-0.004	0.002
Round * Complex	2.97e-05	0.002	0.013	0.990	-0.005	0.005
Round * Message	-0.0013	0.002	-0.537	0.591	-0.006	0.003
Round * Complex * Message	0.0026	0.003	0.781	0.435	-0.004	0.009

Table IV: OLS Regression Results using Full Sample in the Charity Experiment
Notes: The sample considers all guesses all replicators make. Standard Errors are robust to cluster correlation at the replicator level.

Panel A. Rounds 5 to 9 **Panel B.** Rounds 10 to 25

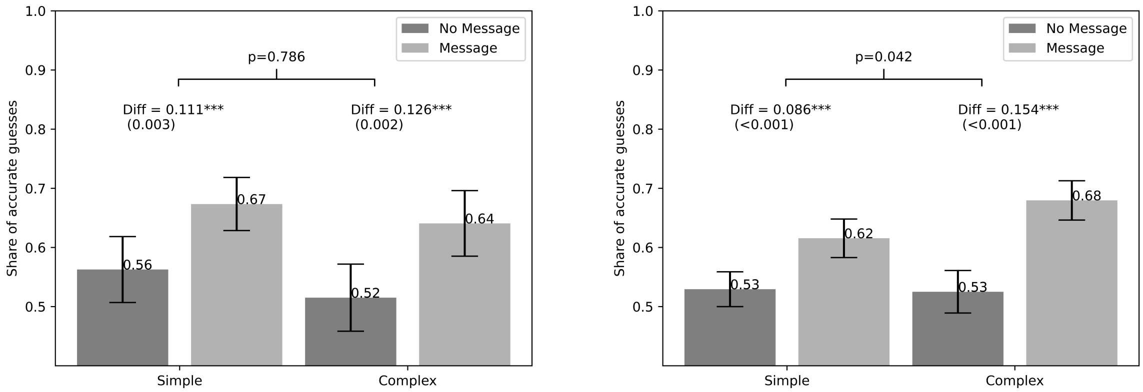


Figure XXIII: Share of perfectly replicable DMs by treatment and condition. Panel A shows accuracy levels in replication with and without message for decision-makers surprised in the first five rounds. Panel B shows the same for those surprised in all other rounds.

Note: Bars show the average likelihood that a replicator guesses a given decision correctly across treatments and conditions. The sample considered in panel A the sample of non-obvious menus for DMs surprised before round 10, and in panel B the sample considers non-obvious menus for DMs surprised in all other rounds; see figure VI in the main body of the paper for the sample that pools all rounds. Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

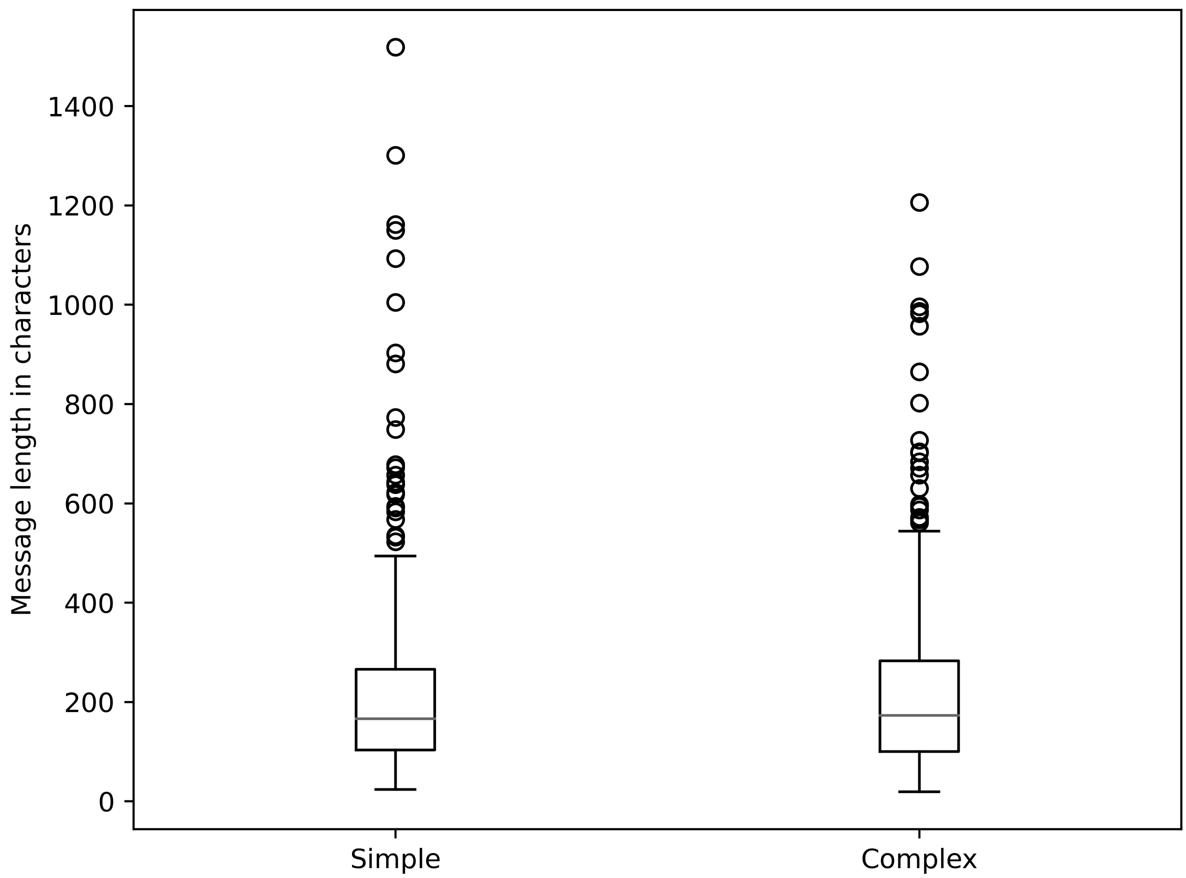


Figure XXIV

Note: Median lengths are 166 and 173 for the simple and complex treatments, respectively.

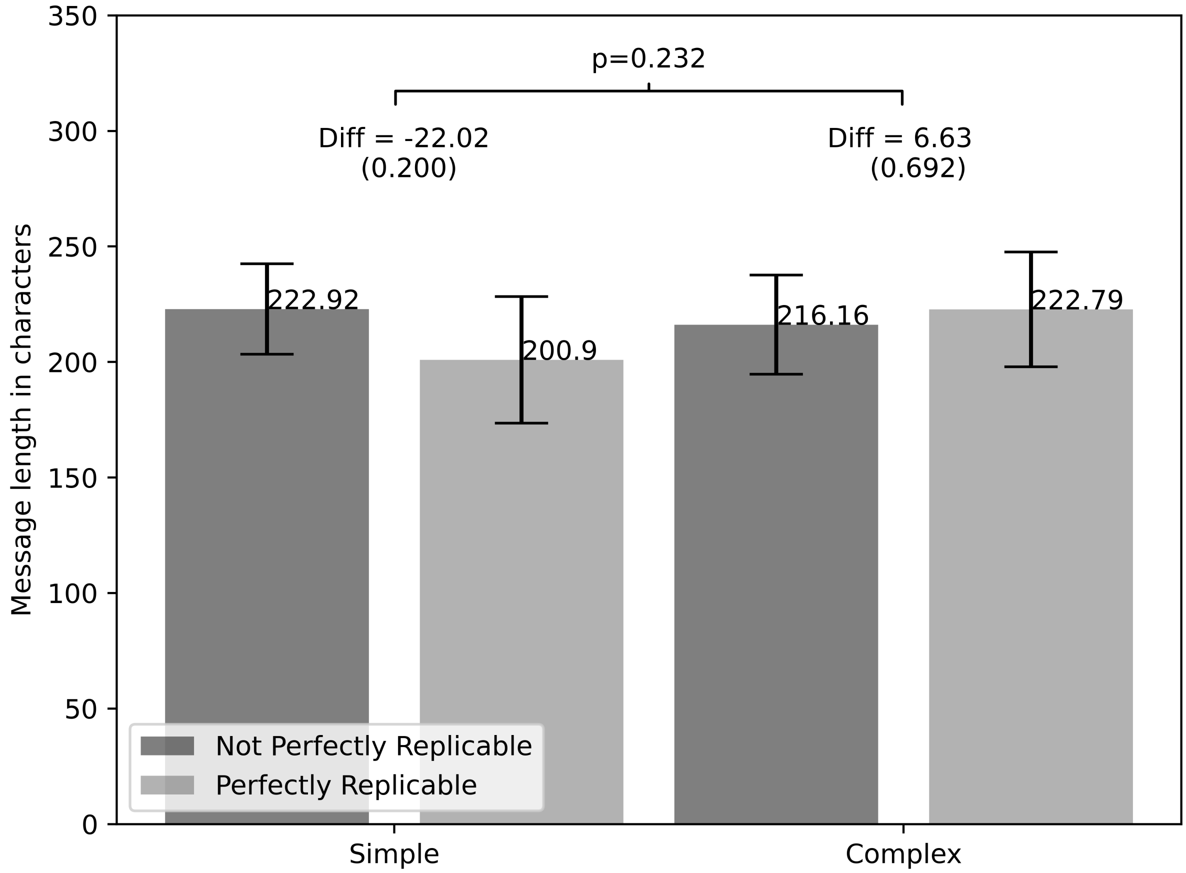


Figure XXV: Message length in characters, for each treatment, split by perfectly and non-perfectly replicable DMs.

<u>Note:</u> Bars show the average length of messages in number of characters for each treatment, split by perfectly and non-perfectly replicable DMs. The sample considered is the sample of all DMs who get matched to a replicator in the no message and message conditions.

Figure XXVI: Self-reports of "Did you find it easy to decide which charity to donate to?"
Note: Bars show the share of DMs who gave each of the responses in the x-axis to the question "Did you find it easy to decide which charity to donate to?" in each treatment. This sample considers all 1000 DMs, and there are also no differences when excluding DMs who saw the Message Task in the first five rounds.

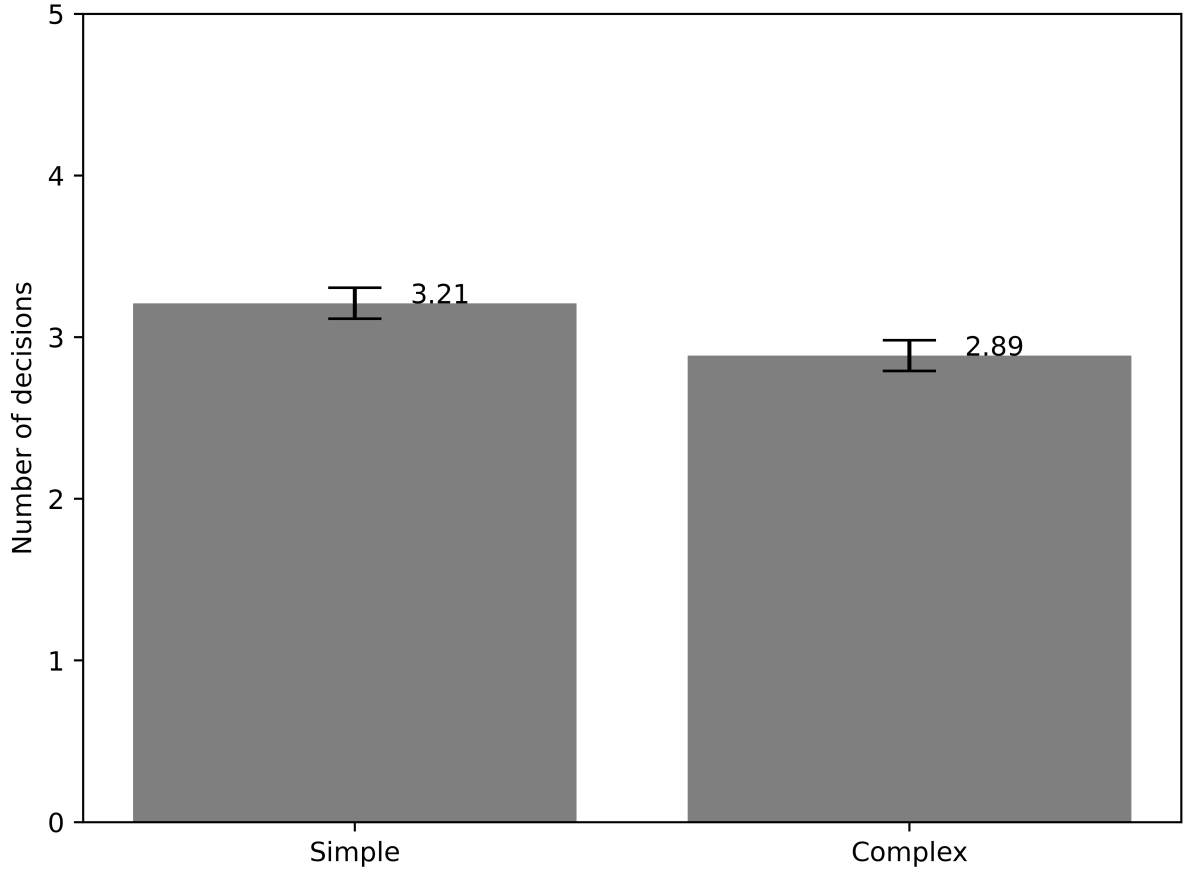


Figure XXVII: DMs responses to: "Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?"

<u>Note:</u> Bars show the share of DMs that give each of the responses in the x-axis. We ask DMs this question once right after writing their message: "Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?" Vertical lines reflect 95% confidence intervals. In parentheses, we show p-values from a t-test at the 95% level.

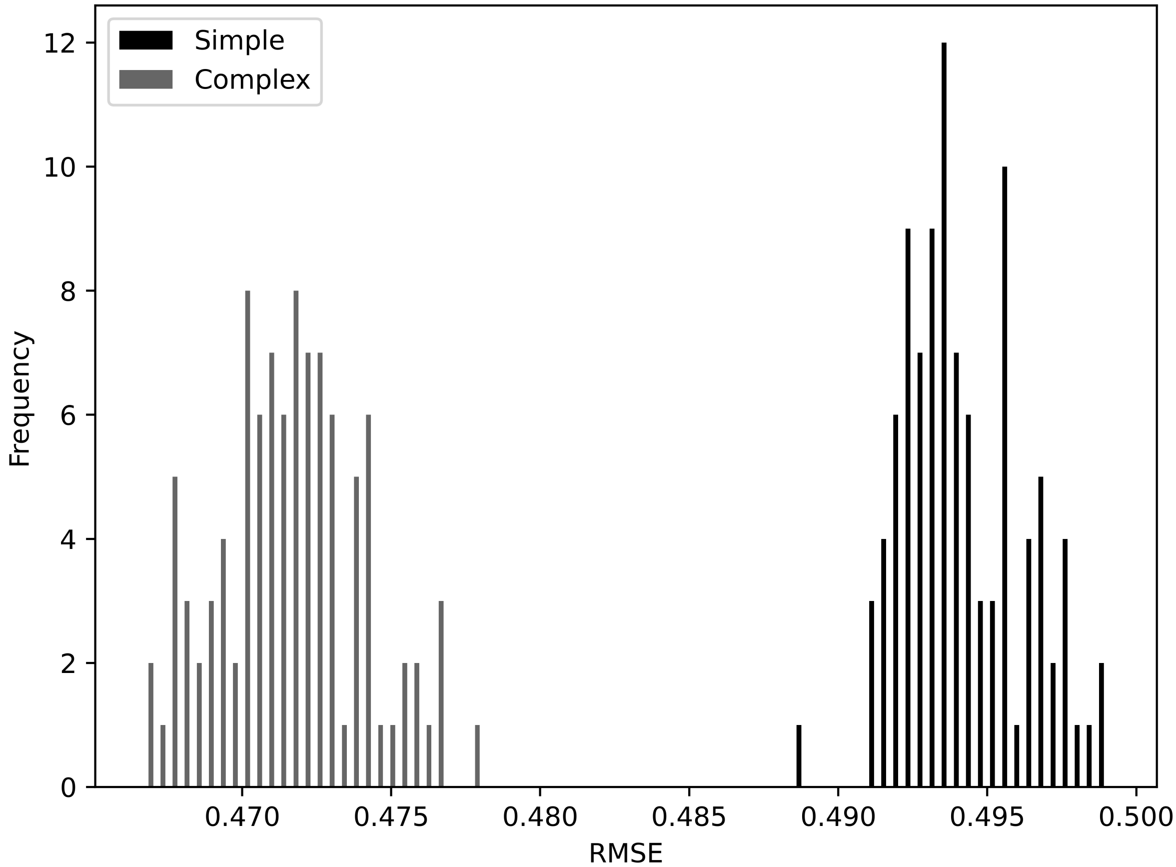


Figure XXVIII: RMSE of Logit Estimation for 100 Splits of the Data in Each Treatment.
Note: Bars show 100 Root Mean Squared Errors for each treatment. These come out of the estimation of a Logit Model as explained in section III.B.1 in the main body of the paper.

A.C. Distribution of differences in moments across treatments

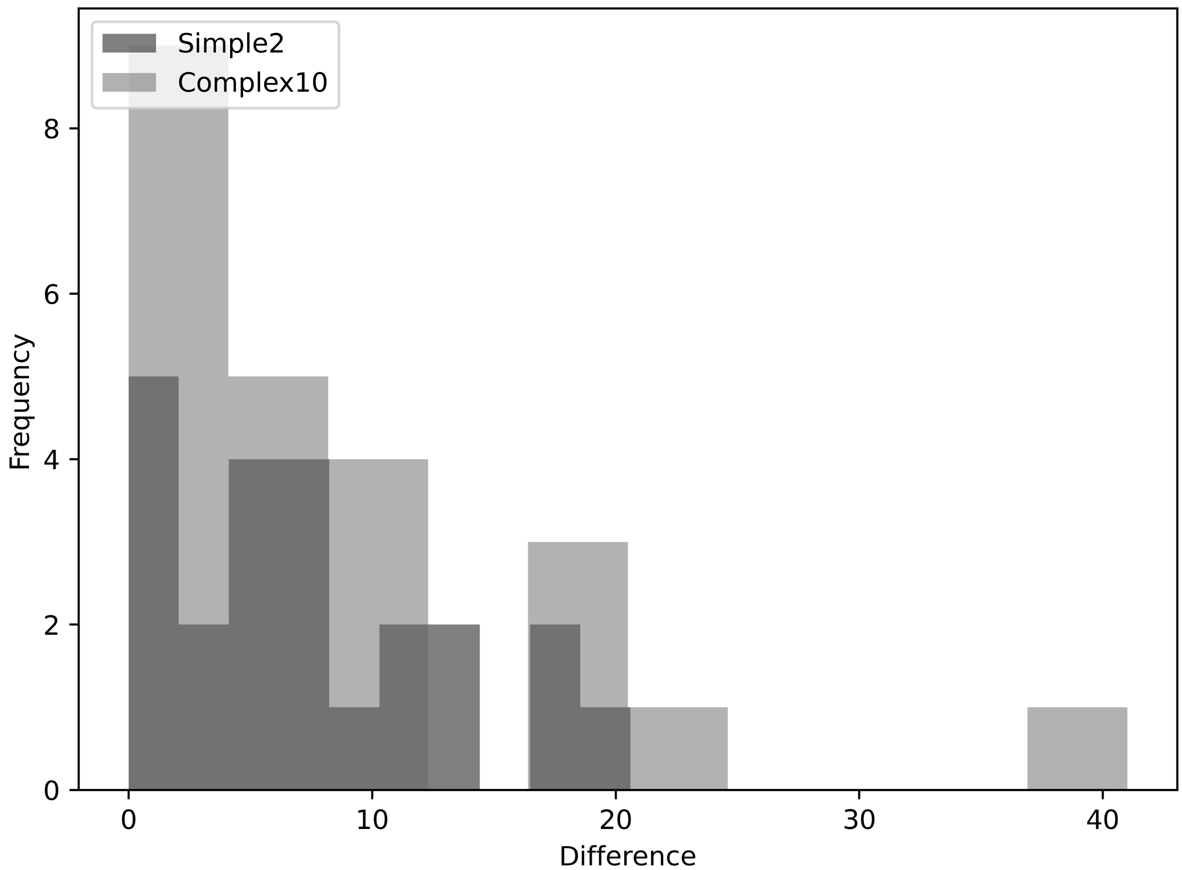


Figure XXIX: Difference in Means

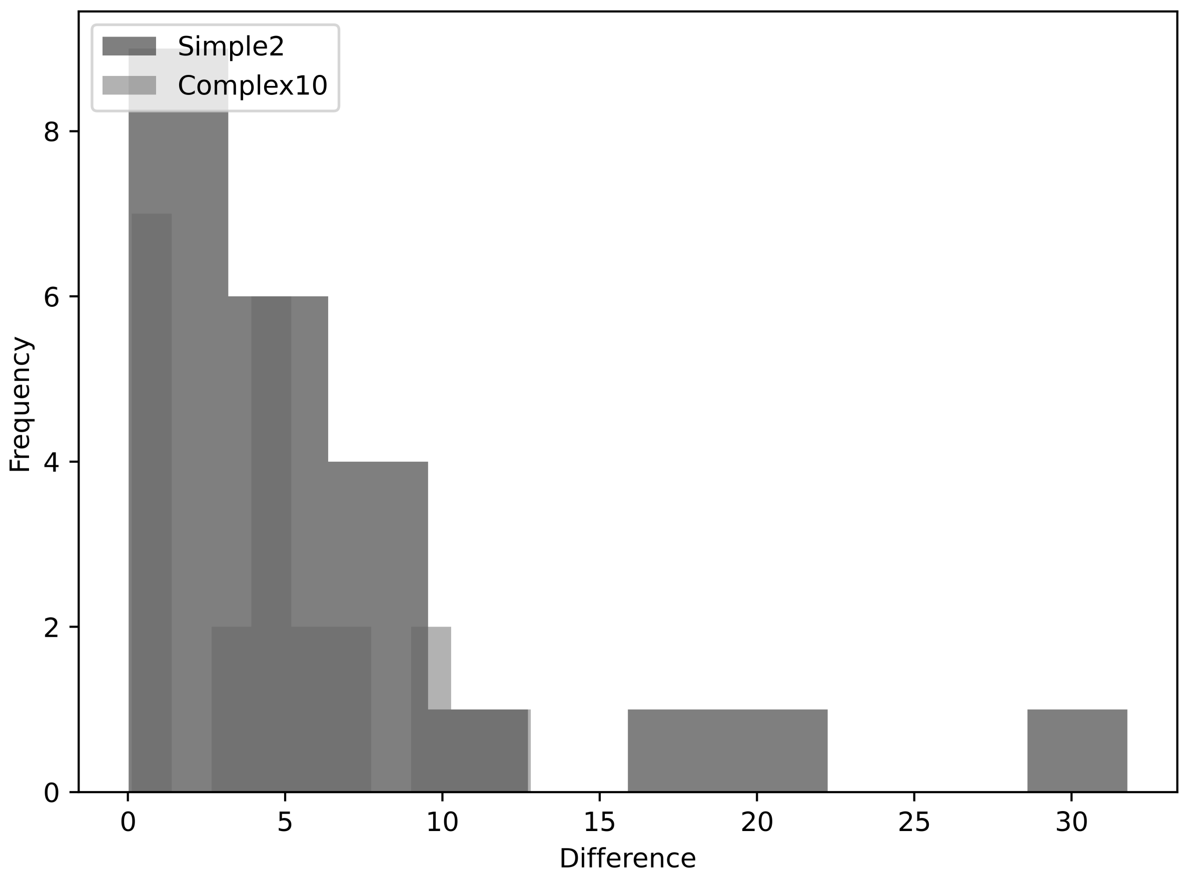


Figure XXX: Difference in Variances

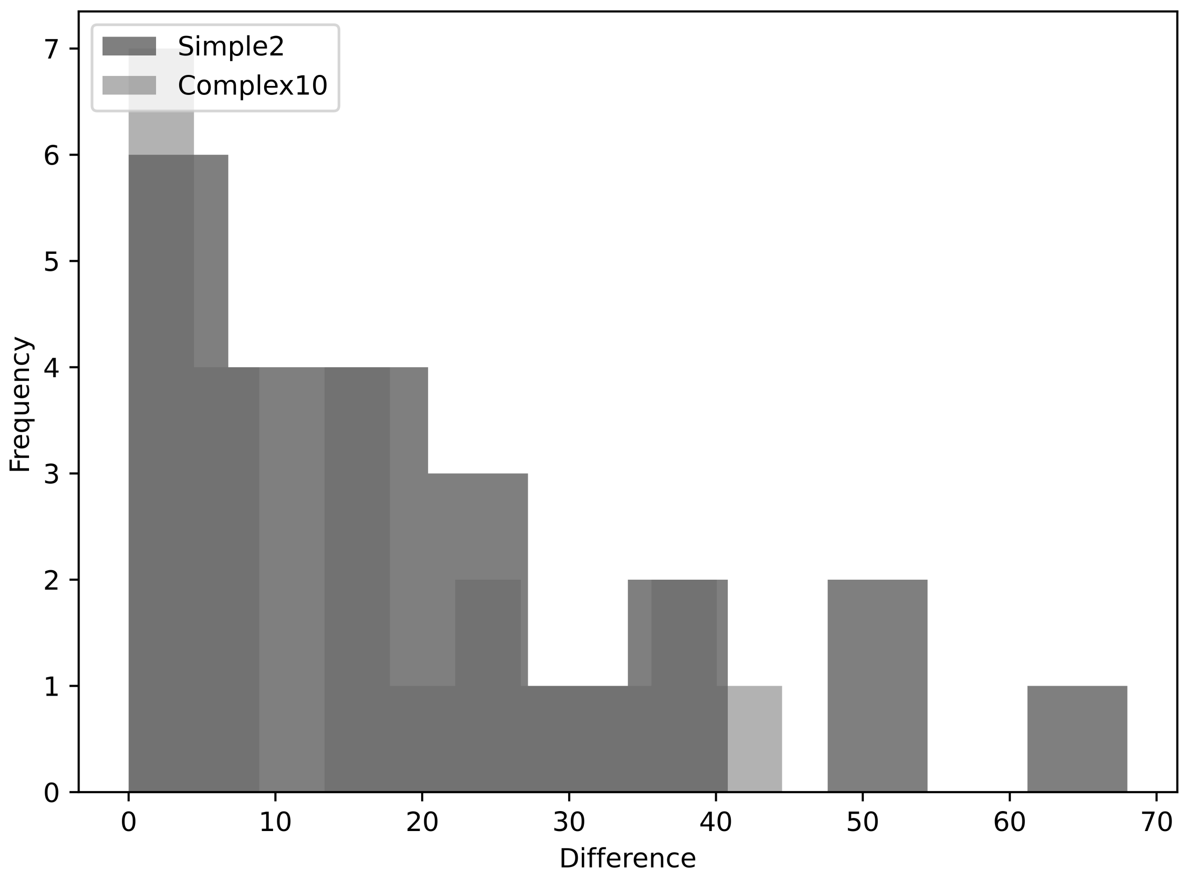


Figure XXXI: Difference in Minimum Outcomes

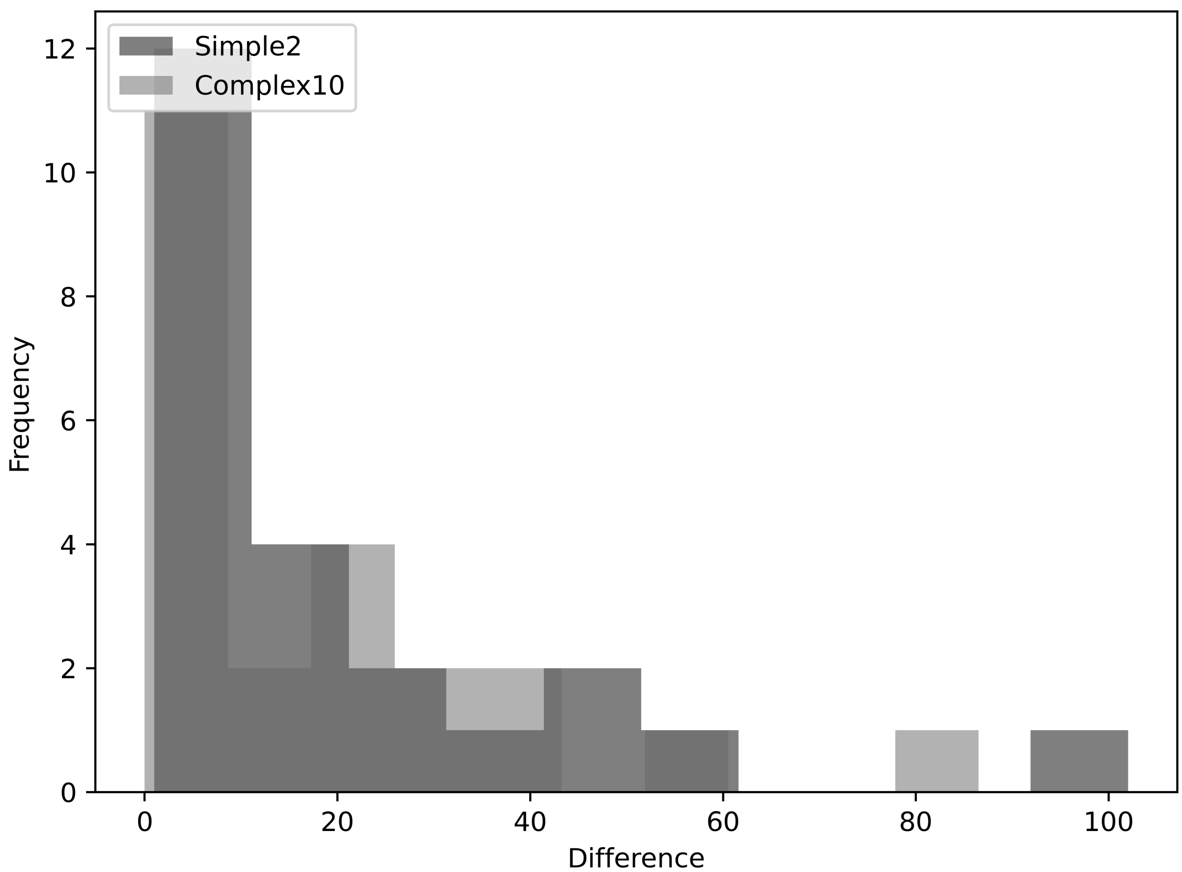


Figure XXXII: Difference in Maximum Outcomes

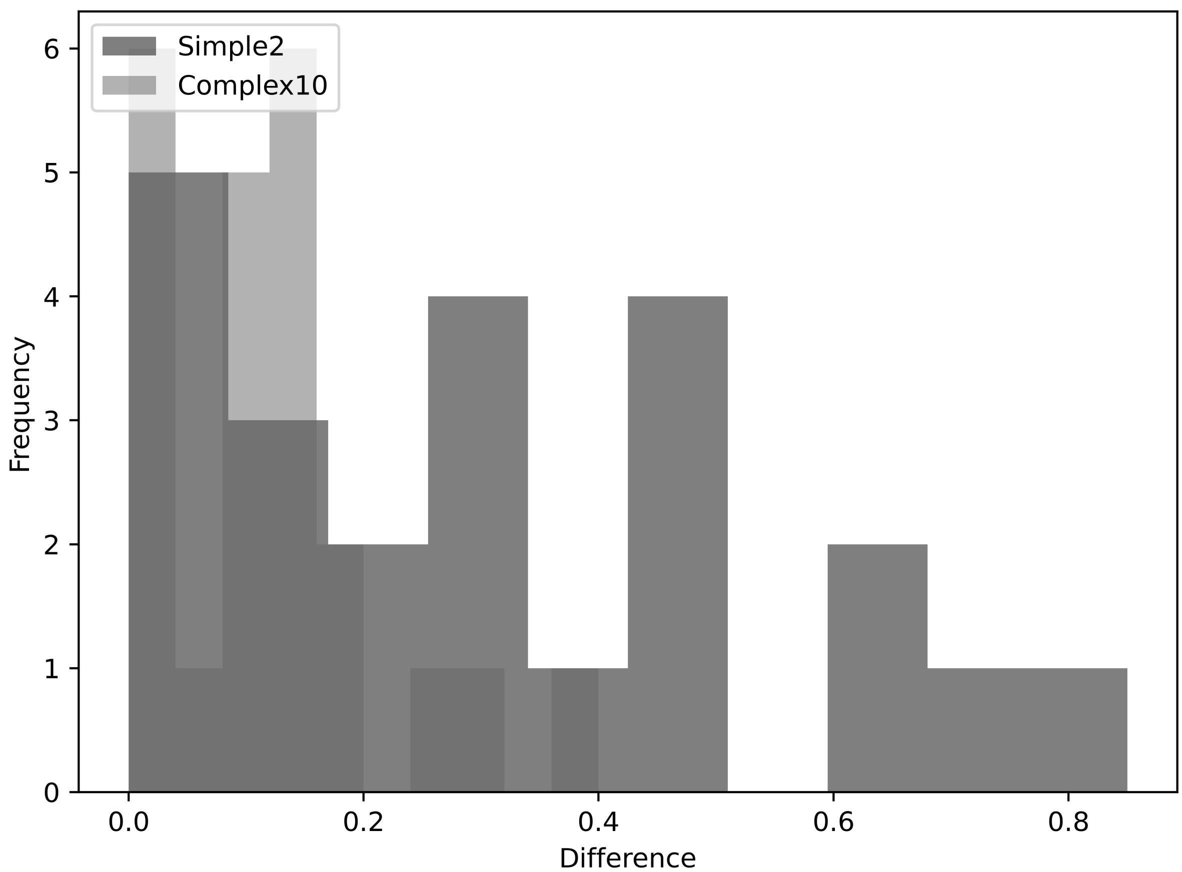


Figure XXXIII: Difference in Chance of Maximum Outcome

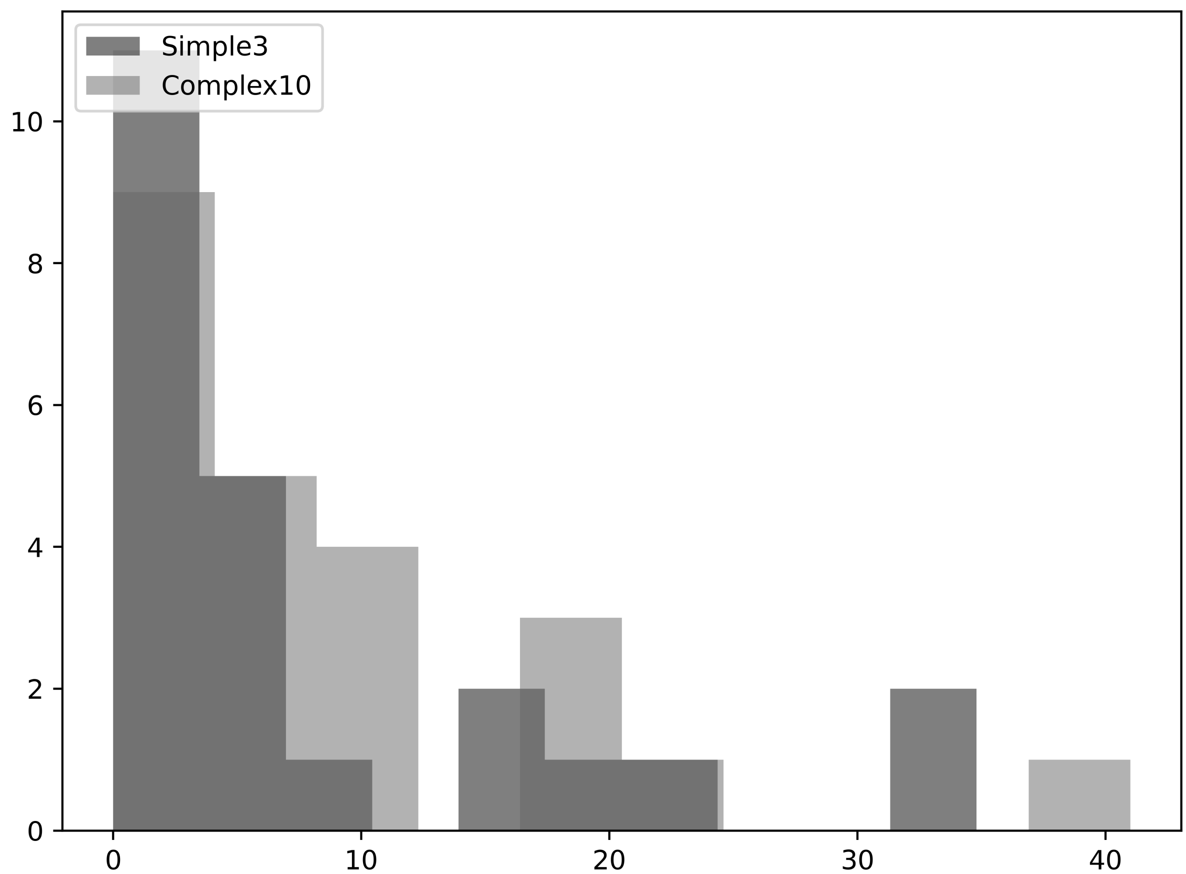


Figure XXXIV: Difference in Means

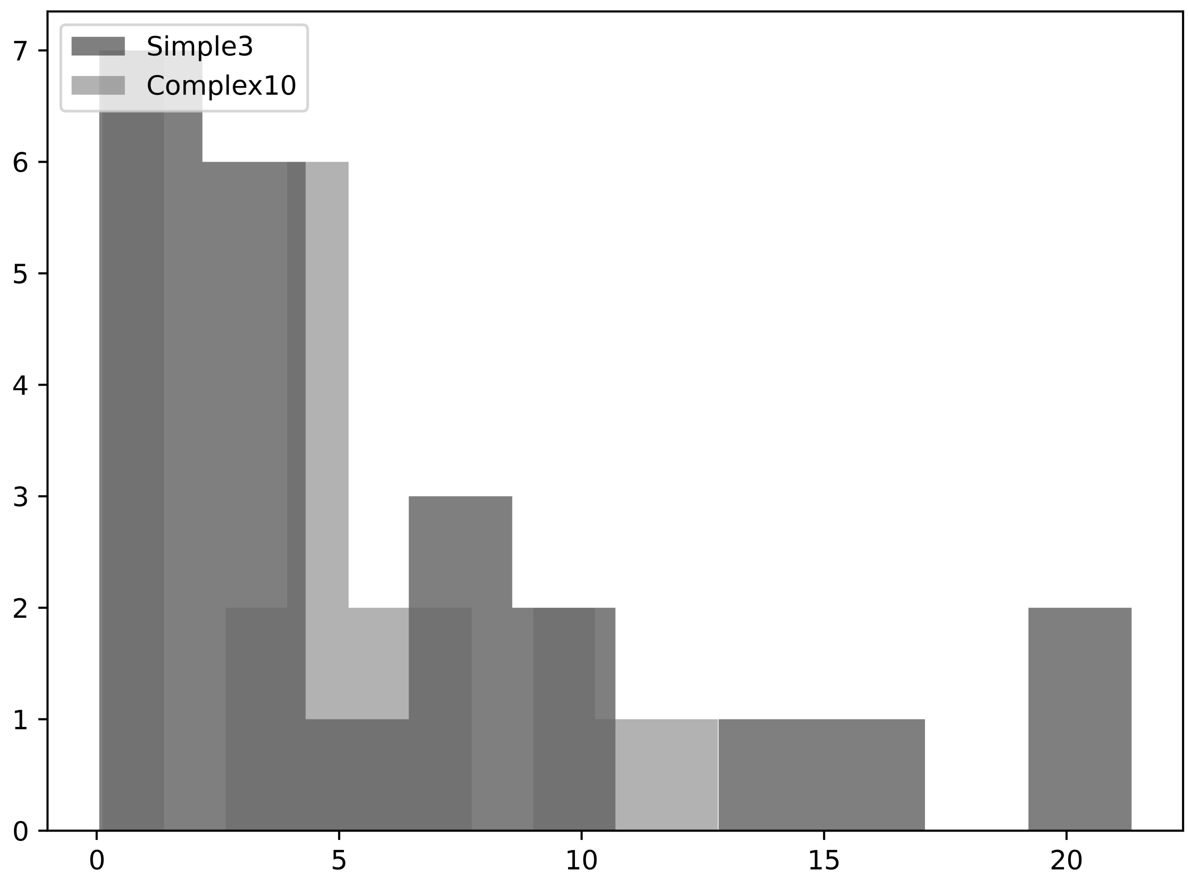


Figure XXXV: Difference in Variances

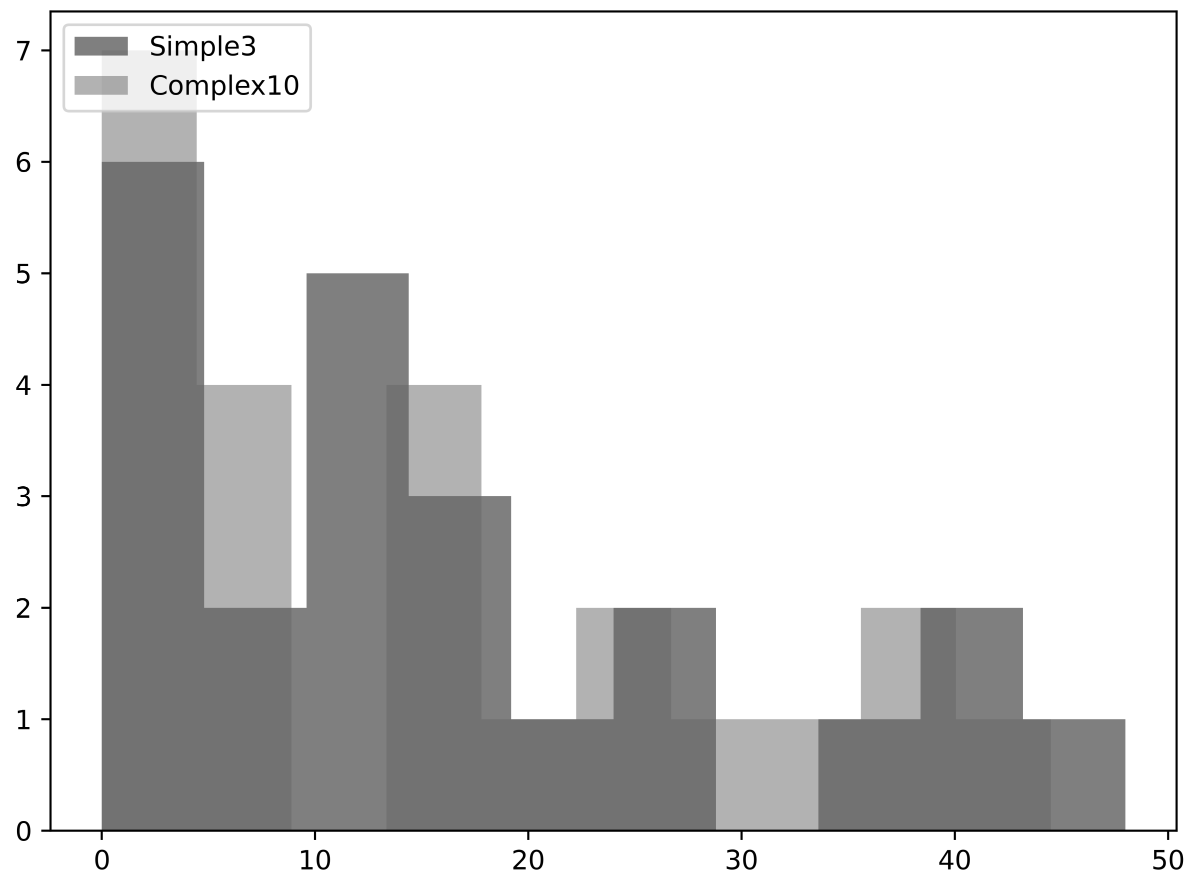


Figure XXXVI: Difference in Minimum Outcomes

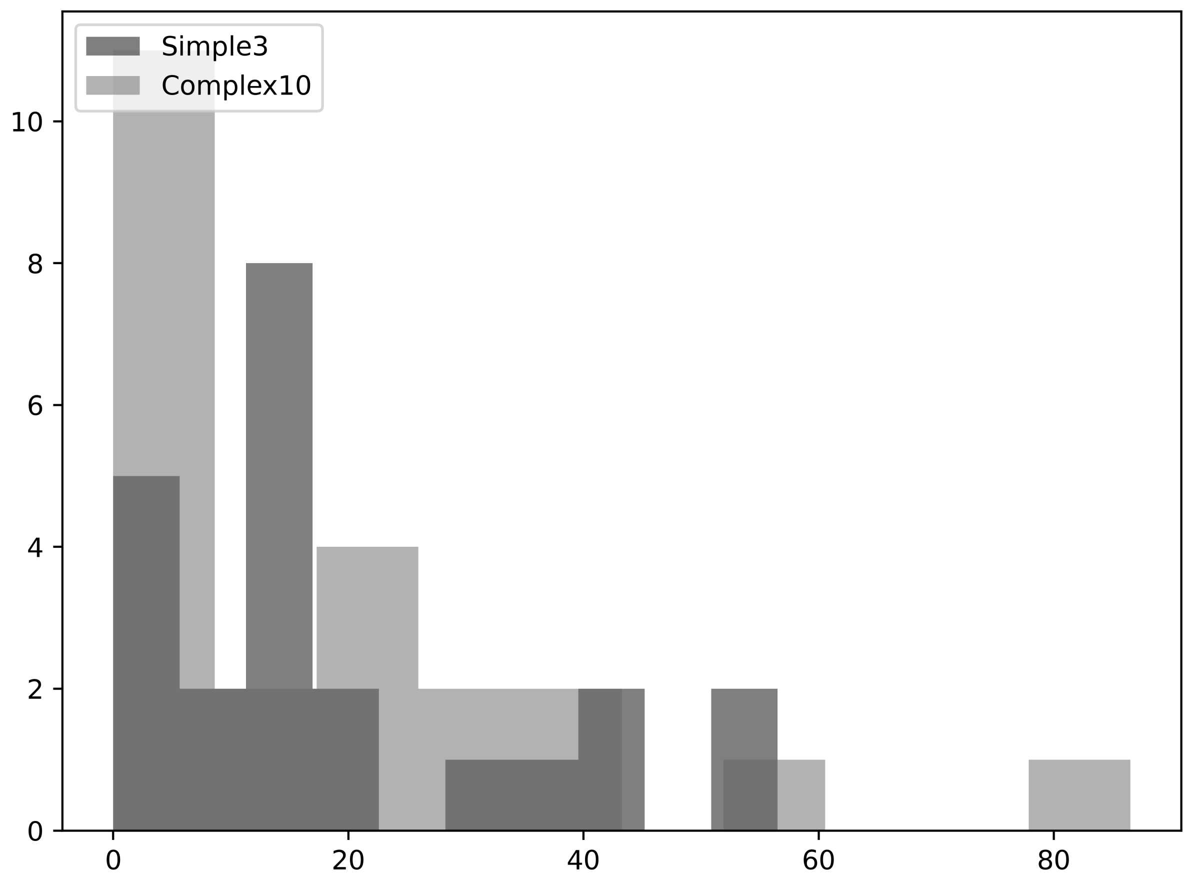


Figure XXXVII: Difference in Maximum Outcomes

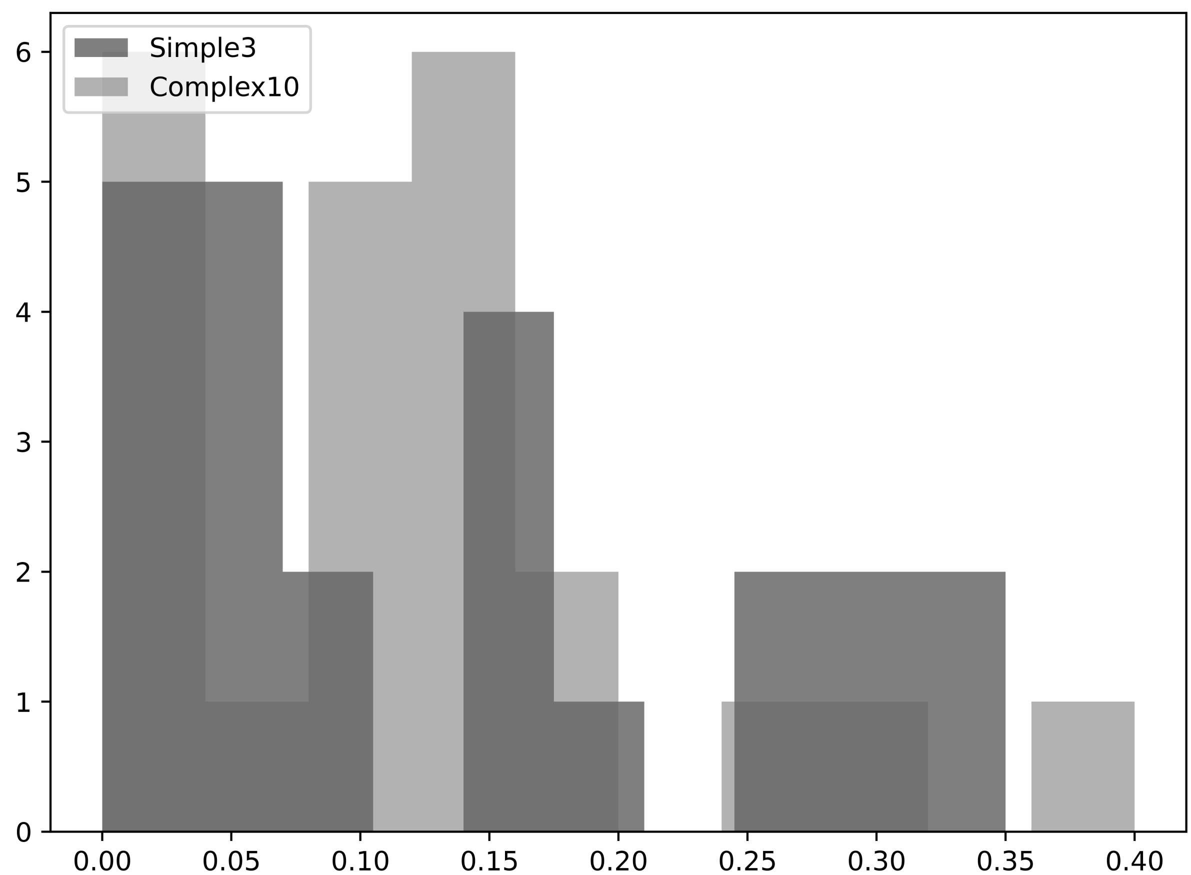


Figure XXXVIII: Caption

Figure XXXIX: Difference in Chance of Maximum Outcome

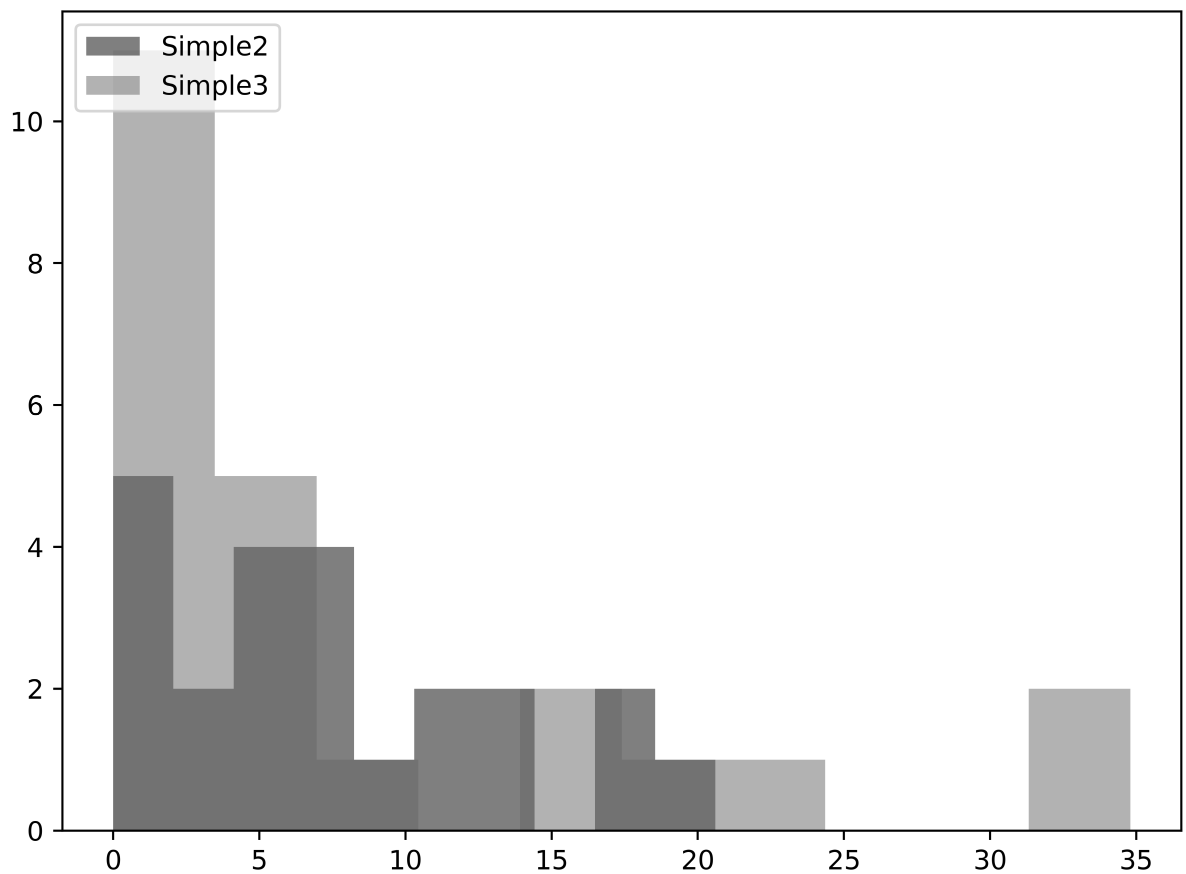


Figure XL: Difference in Means

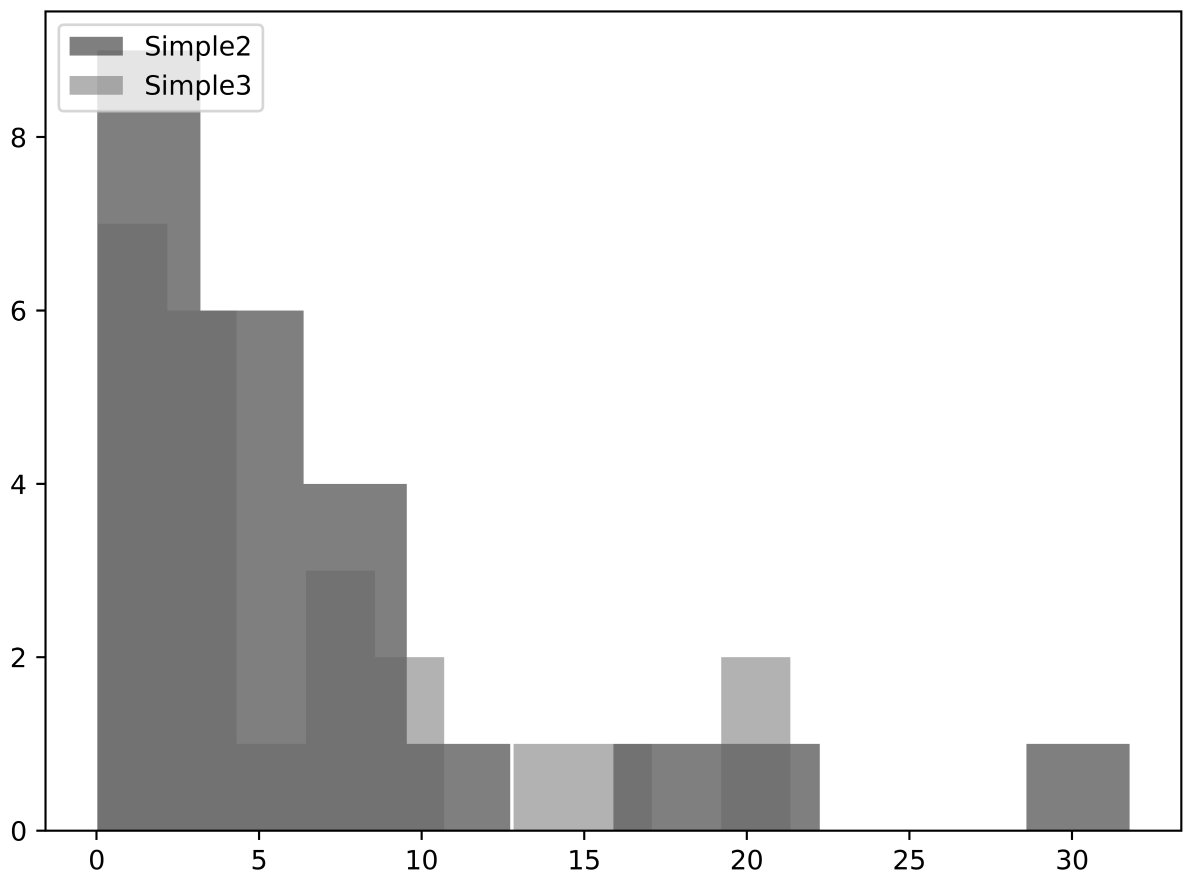


Figure XLI: Difference in Variances

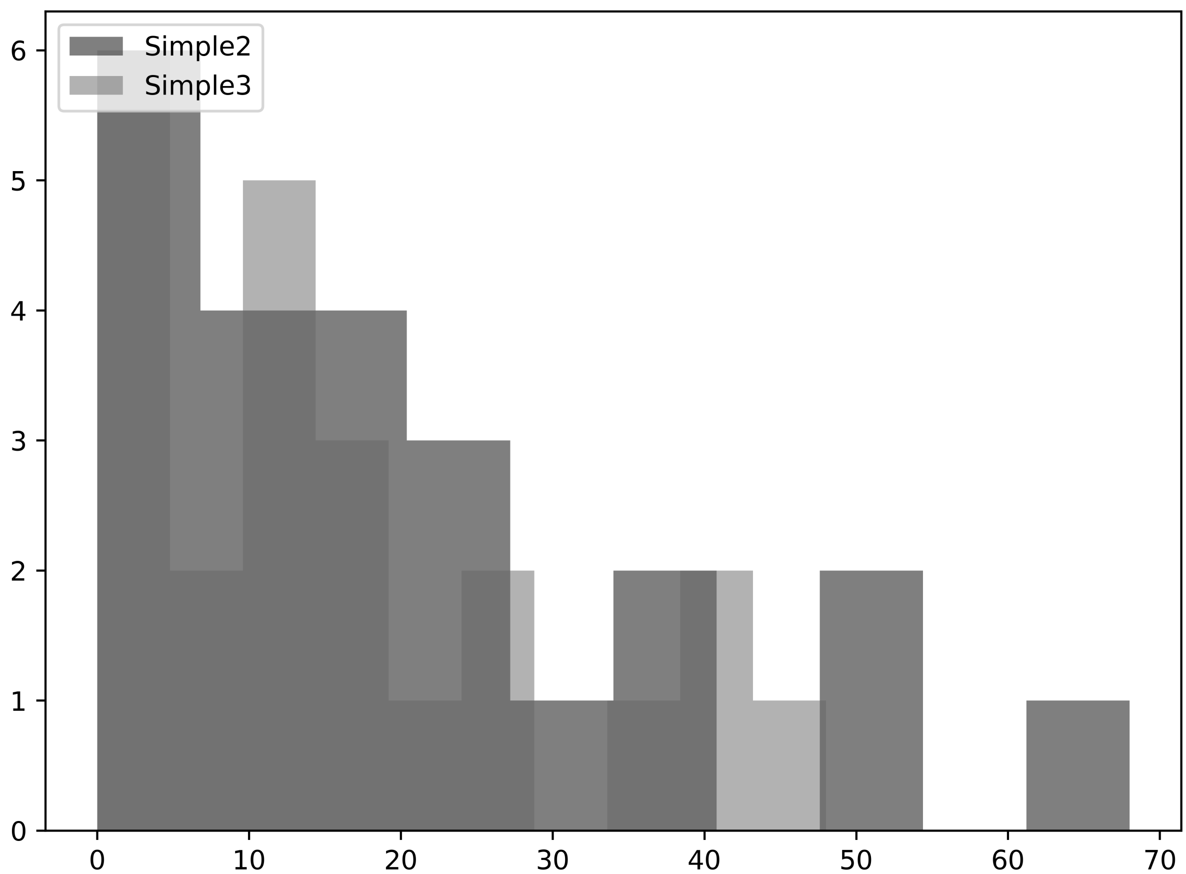


Figure XLII: Difference in Minimum Outcomes

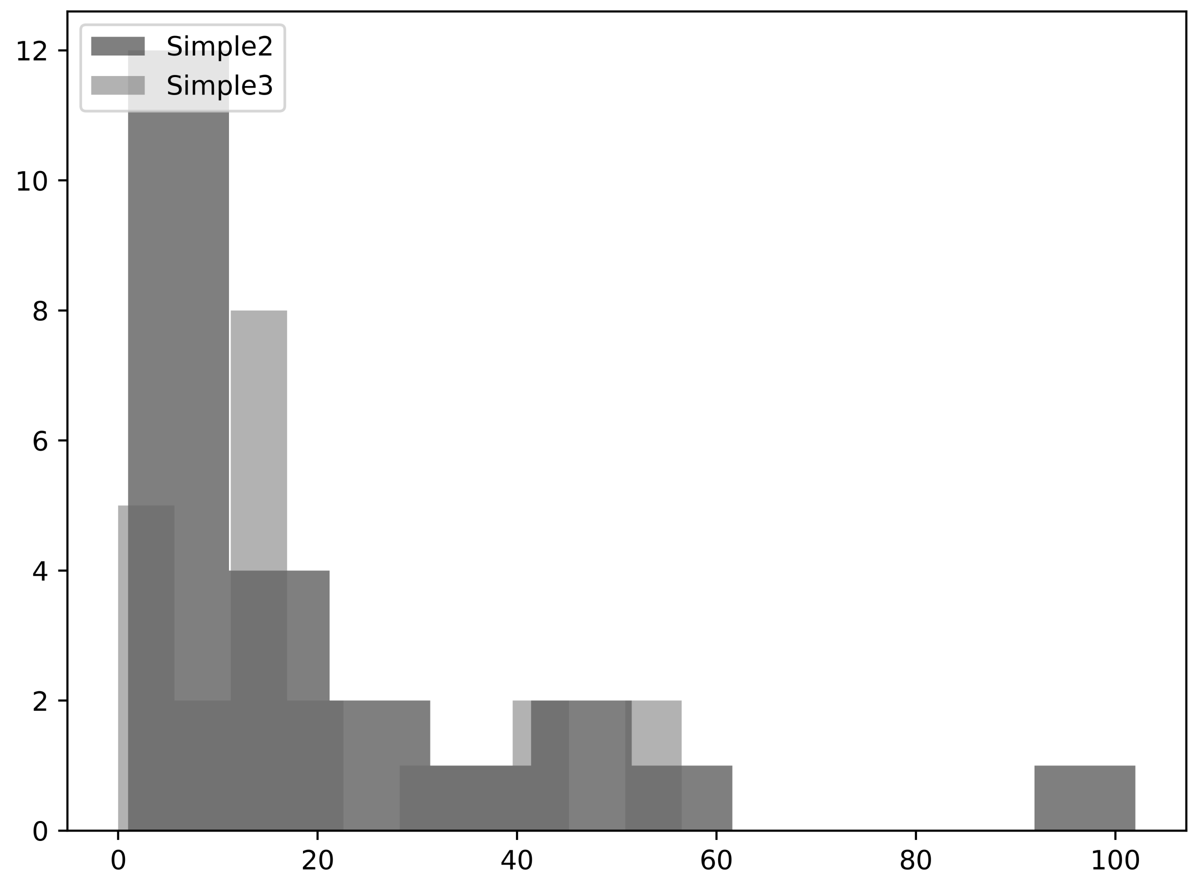


Figure XLIII: Difference in Maximum Outcomes

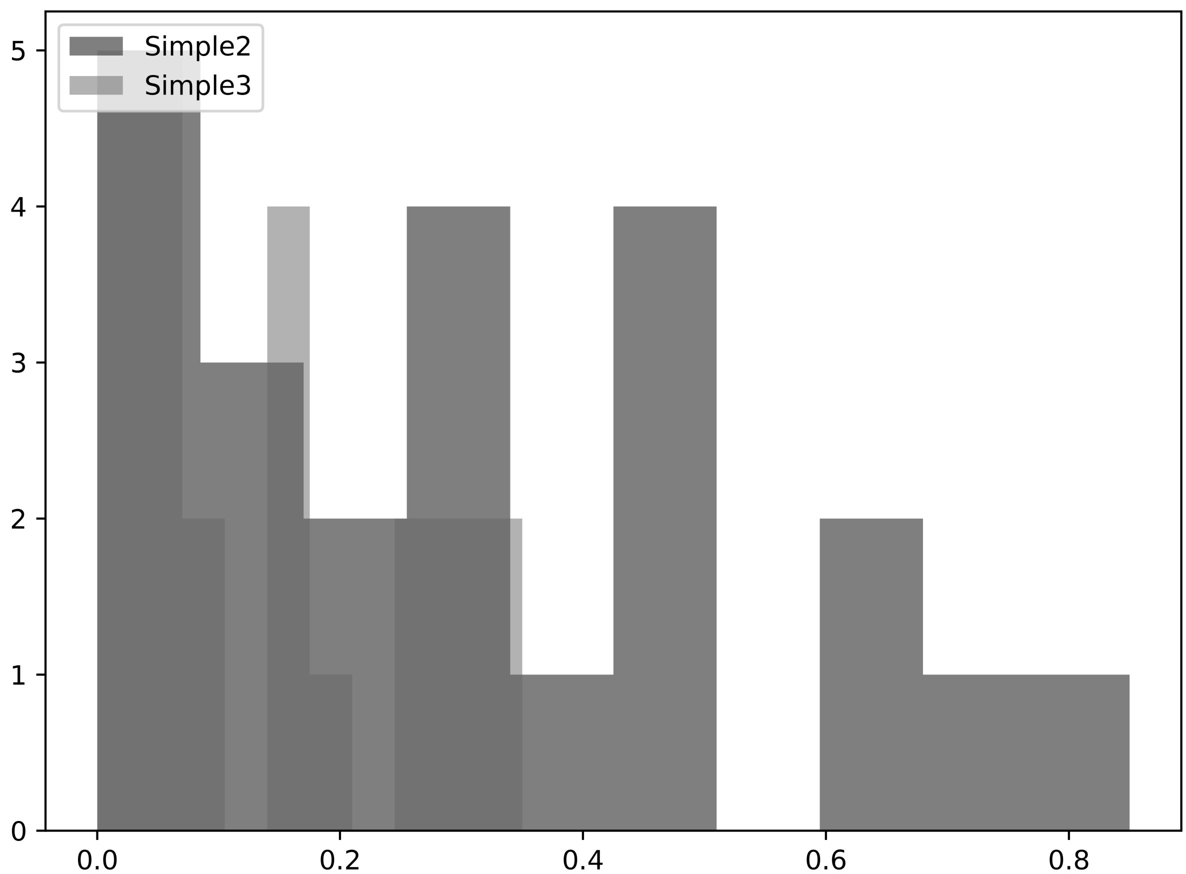


Figure XLIV: Difference in Chance of Maximum Outcome

B. Screenshots

B.A. Experiment 1: Decision-makers' study

Choosing a Lottery

In this study, you will be making 25 decisions. In each decision, we will present you with 2 *lotteries*, and you will choose the lottery you most prefer.

In addition, we will randomly select one of the participants who complete this study. For this participant, we will randomly select one of their 25 decisions, and we will pay them the outcome of the lottery that they selected in this decision.

Thus, please make your decisions thoughtfully and carefully as they could result in a substantial bonus payment of up to $150.

Next

What is a lottery?

A lottery simply specifies the chance of receiving certain payoffs. The chance of each payoff can be anything from 0% to 100%. Each of the lotteries will have 2 possible outcomes.

For example, a lottery looks like the following:

Example Lottery

Outcome	Probability
$2.00	25%
$5.50	75%

You can think of this lottery as paying $2 in 25 out of 100 chances (25% chance), and paying $5.50 in 75 out of 100 chances (75% chance). There are many different possible lotteries.

In each decision, you will see two lotteries on your screen—one on the left, and one on the right.

Your task is simply to choose the lottery you prefer by clicking the associated button. The computer will record your choice and then will present you again with two new lotteries, and so on.

Answer the understanding questions below.

Question 1

Lottery 1

Outcome	Probability
$6.90	5%
$5.65	95%

What is the chance that this lottery pays $6.90?

Question 2

Lottery 2

Outcome	Probability
$1.15	10%
$5.65	90%

Lottery 3

Outcome	Probability
$5.65	45%
$4.55	55%

Which of the two lotteries is more likely to pay exactly $5.65?

When you are ready, continue.

Next

When picking, you will see five additional pieces of information:

- **Average Payment:** (also known as "expected value") This is the average payment the lottery would pay out if it were played many, many times. Lotteries with higher average payments pay more on average, but there is still randomness to how much they pay each time. To calculate the average payment, we multiply each outcome by the probability of that outcome occurring and add this up for all the outcomes.
- **Payment Variability:** (also known as "variance") This is a measure of how much payments can vary. Lotteries with higher payment variability pay amounts that are more spread out, which often means the difference between the larger and lower payments is larger. To calculate the payment variability, we subtract each outcome from the average payment, square it, and then add this up for all the outcomes, weighted by the respective probabilities.
- **Minimum Payment:** This is the minimum possible amount the lottery could pay.
- **Maximum Payment:** This is the maximum possible amount the lottery could pay.
- **Chance of Max Payment:** This is the probability that the lottery pays the Maximum Payment.

We will show you two lotteries, and you can access this additional information about them by clicking on the buttons at the bottom, as we show below. When you click on a button, it will display that information for both lotteries:

Outcome	Probability
$1.15	10%
$5.65	90%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Outcome	Probability
$4.55	55%
$5.65	45%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Please click on the buttons above to get a sense of how this works for the example lotteries above.

When making your 25 choices between lotteries, you will be able to click to learn as many pieces of additional information you want about the lotteries. If you wish, you can use this information to help you make your choices. **Please note that you do not have to use these buttons at all; they are just for your convenience.**

When you are ready, continue.

Next

Choose a Lottery: Round 1

Outcome	Probability
$146.50	5%
$142.00	95%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Outcome	Probability
$107.00	50%
$150.00	50%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Choose Lottery 1

Choose Lottery 2

Remember that we will randomly select a participant and one of their 25 decisions, and pay them the outcome of the lottery that they selected in that decision.

Review Instructions

25 Rounds of Lottery Choices

You've finished your 20th decision

We have an additional task for you before proceeding.

On the next page, **we ask you to *describe* how you were selecting your preferred lotteries.** We call this "your description."

After this study is complete, we will recruit new participants through Prolific, and show *your description* to another Prolific participant. We will then ask this participant to *guess* the lotteries you selected in your past five decisions. You may receive an **additional bonus payment** if this other participant is able to match your choices. Specifically, we will randomly select one of this other participant's guesses. **If the other participant correctly guesses the lottery you picked, you will receive a $5.00 bonus payment.** This other participant will also receive a bonus payment for guessing correctly.

When trying to guess the lottery you picked, the other participant will see your last five decisions in random order. We will renumber the lotteries when they see them. The other participant will see the lotteries and the buttons with additional information, just as you did.

To ensure your understanding, please answer the question below.

How will your message affect your bonus payment?

Continue to message

Write the message to the other participant

Below, please write your message to the other participant describing how you selected lotteries.

Remember, if the other participant is able to guess your decision based on your message, you will earn a bonus payment, so please write your message carefully.

Review Instructions

Please write a message to another participant describing how you made your last five decisions.

This is the only time we will ask you to do this. In the remaining decisions, you only need to select the lottery that you prefer.

When you are ready, click "Submit." You will continue with the remaining lottery choices.

Submit

Guess how the other participant will do in replicating your choices

Please make your decisions carefully. We will randomly pick one participant like yourself to receive $10.00 if their guess is accurate.

Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?

- 0
- 1
- 2
- 3
- 4
- 5

Continue choosing lotteries

Please answer the questions below

Did you describe your decision-making process to the best of your ability?

- ◯ Yes
- ◯ No

If not, why not? Answer in approximately 1-2 sentences.

Do you think you used the same decision-making process in all of the rounds?

- ◯ No, I think it changed
- ◯ Yes, I think I used the same

Would you say that you developed a shortcut to pick a lottery?

- ◯ Yes, I definitely used a shortcut
- ◯ No, I do not think I used any shortcut
- ◯ I sometimes used a shortcut and sometimes did not

Would you say that you developed a rule or procedure to pick a lottery?

- ◯ Yes, I definitely used a rule or procedure
- ◯ No, I do not think I used any rule or procedure
- ◯ I sometimes used a rule or procedure and sometimes did not

How easy was it for you to describe your decision-making process to the other participant?

- ◯ Very easy
- ◯ Somewhat easy
- ◯ Somewhat difficult
- ◯ Very difficult

Why? Answer in approximately 1-2 sentences.

How easy was it for you to decide which lottery to choose?

- ◯ Very easy
- ◯ Somewhat easy
- ◯ Somewhat difficult
- ◯ Very difficult

Why? Answer in approximately 1-2 sentences.

Next

Guess what another participant chose

In this study, your task is to try to *guess* what a previous Prolific participant chose in one of their decisions.

First, we will show you the instructions that the previous participant read before making their decisions. Please read them carefully as understanding them will help you with your task. The other participant's instructions will always be inside a colored box as shown below.

> These instructions are for the other participant.

Continue to read these instructions.

Next

These are the instructions that the other Prolific participant received: Page 1 of 4.

> In this study, you will be making 25 decisions. In each decision, we will present you with 2 *lotteries*, and you will choose the lottery you most prefer.
>
> In addition, we will randomly select one of the participants who complete this study. For this participant, we will randomly select one of their 25 decisions, and we will pay them the outcome of the lottery that they selected in this decision.
>
> **Thus, please make your decisions thoughtfully and carefully as they could result in a substantial bonus payment of up to $150.**

Next

These are the instructions that the other Prolific participant received: Page 2 of 4.

What is a lottery?

A lottery simply specifies the chance of receiving certain payoffs. The chance of each payoff can be anything from 0% to 100%. Each of the lotteries will have 3 possible outcomes.

For example, a lottery looks like the following:

Example Lottery

Outcome	Probability
$2.00	20%
$3.75	50%
$5.50	30%

You can think of this lottery as paying $2 in 20 out of 100 chances (20% chance), paying $3.75 in 50 out of 100 chances (50% chance) and paying $5.50 in 30 out of 100 chances (30% chance). There are many different possible lotteries.

In each decision, you will see two lotteries on your screen—one on the left, and one on the right.

Your task is simply to choose the lottery you prefer by clicking the associated button. The computer will record your choice and then will present you again with two new lotteries, and so on.

The other participant was required to correctly answer the comprehension questions below. To ensure your understanding of the instructions the other participant saw, please answer the questions below.

Question 1

Lottery 1

Outcome	Probability
$5.65	45%
$8.65	50%
$6.90	5%

What is the chance that this lottery pays $6.90?

Question 2

Lottery 2

Outcome	Probability
$5.90	5%
$1.15	10%
$5.65	85%

Lottery 3

Outcome	Probability
$5.65	40%
$8.65	35%
$4.55	25%

Which of the two lotteries is more likely to pay exactly $5.65?

When you are ready, continue.

Next

These are the instructions that the other Prolific participant received: Page 3 of 4.

When picking, you will see five additional pieces of information:

- **Average Payment:** (also known as "expected value") This is the average payment the lottery would pay out if it were played many, many times. Lotteries with higher average payments pay more on average, but there is still randomness to how much they pay each time. To calculate the average payment, we multiply each outcome by the probability of that outcome occurring and add this up for all the outcomes.
- **Payment Variability:** (also known as "variance") This is a measure of how much payments can vary. Lotteries with higher payment variability pay amounts that are more spread out, which often means the difference between the larger and lower payments is larger. To calculate the payment variability, we subtract each outcome from the average payment, square it, and then add this up for all the outcomes, weighted by the respective probabilities.
- **Minimum Payment:** This is the minimum possible amount the lottery could pay.
- **Maximum Payment:** This is the maximum possible amount the lottery could pay.
- **Chance of Max Payment:** This is the probability that the lottery pays the Maximum Payment.

We will show you two lotteries, and you can access this additional information about them by clicking on the buttons at the bottom, as we show below. When you click on a button, it will display that information for both lotteries:

Outcome	Probability
$5.65	85%
$5.90	5%
$1.15	10%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Outcome	Probability
$8.65	35%
$5.65	40%
$4.55	25%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Please click on the buttons above to get a sense of how this works for the example lotteries above.

When making your 25 choices between lotteries, you will be able to click to learn as many pieces of additional information you want about the lotteries. If you wish, you can use this information to help you make your choices. **Please note that you do not have to use these buttons at all; they are just for your convenience.**

When you are ready, continue.

Next

After reading these instructions and correctly answering understanding questions about them, the previous participant then went on to make 25 decisions. At a random point throughout these 25 decisions, they received the additional instructions below. <u>Please read this very carefully, as this directly determines your task today.</u>

These are the instructions that the other Prolific participant received: Page 4 of 4

We have an additional task for you before proceeding.

On the next page, **we ask you to *describe* how you were selecting your preferred lotteries.** We call this "your description."

After this study is complete, we will recruit new participants through Prolific, and show *your description* to another Prolific participant. We will then ask this participant to *guess* the lotteries you selected in your past five decisions. You may receive an **additional bonus payment** if this other participant is able to match your choices. Specifically, we will randomly select one of this other participant's guesses. **If the other participant correctly guesses the lottery you picked, you will receive a $5.00 bonus payment.** This other participant will also receive a bonus payment for guessing correctly.

When trying to guess the lottery you picked, the other participant will see your last five decisions in random order. We will renumber the lotteries when they see them. The other participant will see the lotteries and the buttons with additional information, just as you did.

The other participant wrote a message here.

To ensure your understanding of the instructions the other participant saw, please answer the questions below.

When the other participant was writing their message, they knew that they would receive a bonus payment if:

When you are ready, click "Submit."

Submit

Only for participants in the message condition

Your task in this study: Guess 5 choices for each of 3 participants

Your task is to try to replicate previous participants' choices based on their message that describes their decision-making process. We will show you the description they wrote of their decision-making process. We will show you the decisions of three participants. For each of the three participants, you will see 5 decisions. In each, you will see the exact two lotteries the other participant faced. Your task is to select the lottery you think the previous participant chose, given their description of their decision-making process.

We will randomly select one of these decisions. If you correctly guess the chosen lottery, you will earn a $5.00 bonus, and the previous participant will also earn a $5.00 bonus. If you guess incorrectly, neither of you will earn a bonus.

To ensure your understanding of the instructions the other participant saw, please answer the questions below.

How should you make your decisions in order to maximize your bonus payment?

Continue to start your task.

Start Task

Guess what Participant 1 chose

Participant 1 wrote this message:

> *i picked the lotteries with the best chance of getting alot of money*

Participant 1: Decision 1 out of 5.

Below, please select the lottery that you think Participant 1 chose in this decision given how they described their decision-making process in the message above.

Outcome	Probability
$20.00	20%
$111.00	60%
$55.50	20%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Outcome	Probability
$50.00	20%
$110.00	60%
$10.00	20%

Additional information:

Average Payment | Payment Variability | Minimum Payment | Maximum Payment | Chance of Max Payment

Choose Lottery 1

Choose Lottery 2

Review Instructions

Each participant sees 5 pages like this for each of 3 DMs, for a total of 15. Participants in the No Message condition do not see the message in the blue box.

You are done with Participant 1

Out of your 5 decisions, how many do you think you guessed correctly? We will randomly pick one participant like yourself to receive $10 if their guess is accurate.

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5

Did you find the other participant's message comprehensible?

- ○ Yes, very comprehensible
- ○ Somewhat comprehensible
- ○ No, I found it generally hard to comprehend

How easy or hard did you find it to guess the chosen lottery based on the message?

- ○ Very easy
- ○ Easy
- ○ Neither easy nor hard
- ○ Hard
- ○ Very hard

Did the message feel like a step-by-step (or single step) process?

- ○ Yes, very much so
- ○ Somewhat
- ○ Not really

Continue to see the next participant, who will have a different message .

Next

B.C. Experiment 2: Decision-makers' study

Donation to a Charity

In this study, you will be making 25 decisions. In each decision, we will present you with the descriptions of **6 *charities***, and you will choose the charity you'd most prefer to make a donation to.

In addition, we will randomly select one of the participants who complete this study. For this participant, we will randomly select one of their 25 decisions, and we will make a $1,000 donation to the charity that they selected in this decision.

Thus, please make your decisions thoughtfully and carefully as they could result in a substantial donation to the charity you select.

Next

Charity Information

While we do not include the names of these charities, we include relevant information on the charity's cause and the charity's efficiency.

Here is an example charity:

Area of work: Youth Education Programs and Services
Location: Los Angeles, CA
Program Expense Ratio: 93.15%
Administrative Expenses: 15.20%
Fundraising Expenses: 7.50%
Fundraising Efficiency: $0.12
Working Capital Ratio: 0.07 years
Program Expense Growth: 12.57%
Liabilities to Assets: 1.50%

This information is collected on each charity by third-party charity evaluators. The purpose of many of these measures is to assess quality and efficiency of the charity. Below, we give you the definitions of these measures. Please read them carefully.

Area of work: Type of charity

Location: Location of the charity's operating headquarters

Program expense ratio: The fraction of the charity's total expenses (administrative expenses + fundraising expenses + program expenses) that go toward the program rather than toward administrative and fundraising expenses. This measure reflects the percent of the charity's total expenses that it spends on the programs and services it exists to deliver. A charity with a 100% program expense ratio spends *all* of its money on its charitable mission, while a charity with a 0% program expense ratio spends *none* of its money on its charitable mission.

Administrative expenses: The percent of the charity's total budget that it spends on overhead, administrative staff and associated costs, and organizational meetings. A charity with 100% administrative expenses spends *all* of its money on overhead and other administrative expenses, while a charity with 0% administrative expenses spends *none* of its money on overhead and other administrative expenses.

Fundraising expenses: The percent of the charity's total budget that it spends on fundraising, including campaign printing, publicity, mailing, etc. A charity with 100% fundraising expenses spends *all* of its money on fundraising, while a charity with 0% fundraising expenses spends *none* of its money on fundraising.

Fundraising efficiency: The amount that the charity spends in order to raise $1 in charitable contributions.

Working capital ratio: How long a charity could sustain its level of spending using its net available assets. A charity with a working capital ratio of 0 would shut down immediately without additional funds, while a charity with a working capital ratio of 20 years could sustain its current level of spending for 20 more years without additional funds.

Program expense growth: The average annual growth of program expenses. A charity with a 100% program expense growth is spending twice as much each year as it did the year before, while a charity with a -50% program expense growth is spending half as much each year as it did the year before.

Liabilities to assets: The charity's total liabilities divided by their total assets. Charities, like other organizations, need to be mindful of their liabilities in relation to their assets. This ratio is an indicator of an organization's solvency and/or long-term sustainability, with a lower ratio indicating larger long-term sustainability.

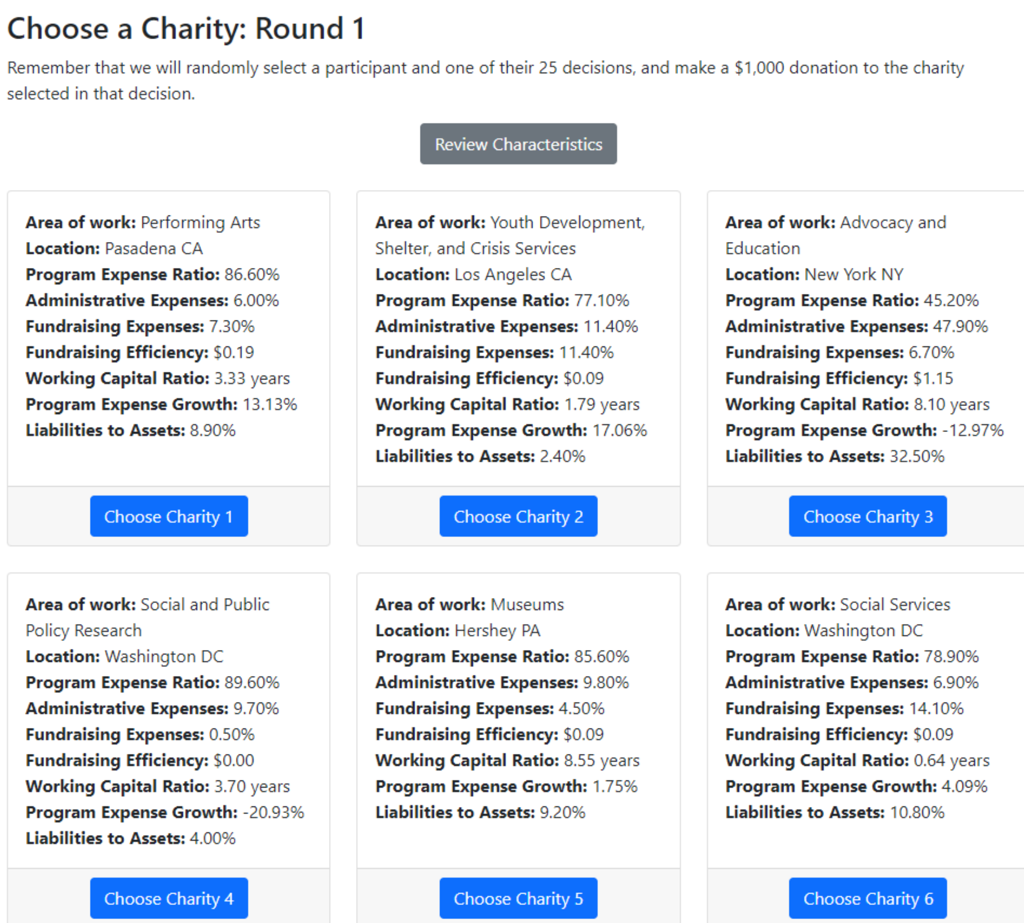
Imagine a charity that donates books to children in need. This charity spends $100 each day. $5 of this goes to fundraising and $65 goes to their CEO. The remaining $30 goes toward books for the children. What is this charity's program expense ratio?

Imagine a charity that just raised $365,000. This charity spends $1,000 per day. What is this charity's working capital ratio?

Imagine two charities: One charity has a fundraising efficiency of $0.20 and the other has a fundraising efficiency of $0.75. Both charities spend $100 on fundraising materials. Which charity would raise more money from this fundraising?

When you are ready, continue.

Next



25 Rounds of Charity Choices

You've finished your 5th decision

We have an additional task for you before proceeding.

On the next page, **we ask you to *describe* how you were selecting your preferred charities.** We call this "your description."

After this study is complete, we will recruit new participants through Prolific, and show *your description* to another Prolific participant. We will then ask this participant to *guess* the charities you selected in your past five decisions. You will receive an **additional bonus payment** if this other participant is able to match your choices. Specifically, we will randomly select one of the guesses that this other participant makes. **If this guess correctly picks the charity you chose, then you will receive a $5.00 bonus.** This other participant will also receive a bonus payment for guessing correctly.

When trying to guess the charity you picked, the other participant will see your last five decisions in random order. Moreover, they will only see two charities instead of all six: they will see the one you picked alongside one other randomly selected charity from that decision. We will renumber the charities when they see them.

To ensure your understanding, please answer the question below.

How will your message affect your bonus payment?

Continue to message

Write the message to the other participant

Below, please write your message to the other participant describing how you selected charities. You can click to review the instructions for writing your message and the list of charity characteristics. We have also provided a list of all the areas of work that charities could be classified into in case you want to reference them in your message.

Remember, if in a randomly picked guess the other participant guesses your decision based on your message, you will earn a bonus payment, so please write your message carefully.

Review Instructions | Review Characteristics | Review Areas of Work

Please write a message to another participant describing how you made your last five decisions.

This is the only time we will ask you to do this. In the remaining decisions, you only need to select the charity you would rather donate to.

When you are ready, click "Submit." You will continue with the remaining charity choices.

Submit

Guess how the other participant will do in replicating your choices

Please make your decisions carefully. We will randomly pick one participant like yourself to receive $10.00 if their guess is accurate.

Out of your 5 decisions, how many do you think the other participant will be able to guess correctly based on your description?

- 0
- 1
- 2
- 3
- 4
- 5

Continue choosing charities

Please answer the questions below

In the message you wrote, did you describe your decision-making process to the best of your ability?

- ◯ Yes ◯ No

If not, why not? Please describe your experience describing your decision-making process.

Do you think your decision-making process changed throughout the 25 rounds?

- ◯ Yes, it changed a lot
- ◯ Yes, it somewhat changed
- ◯ No, it did not change

Were you always considering all of the charities and all of the information about them?

- ◯ I was always considering all of the charities but ignored some information about them
- ◯ I always considered all information about the charities but ignored some charities entirely
- ◯ I always considered all of the charities and all of their information
- ◯ I ignored some charities entirely and ignored some information about the others

Would you say that you developed a shortcut to pick a charity?

- ◯ Yes, I definitely used shortcuts
- ◯ No, I do not think I used any shortcut
- ◯ I sometimes used shortcuts and sometimes did not

Select the charity characteristic that you found most important when choosing

- ◯ Area of Work
- ◯ Location
- ◯ Program Expense Ratio
- ◯ Administrative Expenses
- ◯ Fundraising Expenses
- ◯ Fundraising Efficiency
- ◯ Working Capital Ratio
- ◯ Program Expense Growth
- ◯ Liabilities to Assets

Select the charity characteristic that you found second most important when choosing

- ◯ Area of Work
- ◯ Location
- ◯ Program Expense Ratio
- ◯ Administrative Expenses
- ◯ Fundraising Expenses
- ◯ Fundraising Efficiency
- ◯ Working Capital Ratio
- ◯ Program Expense Growth
- ◯ Liabilities to Assets

Did you find it easy to describe your decision-making process to the other participant?

- ◯ Yes, very easy
- ◯ Yes, somewhat easy
- ◯ No, somewhat hard
- ◯ No, very hard

Did you find it easy to decide which charity to donate to?

- ◯ Yes, very easy
- ◯ Yes, somewhat easy
- ◯ No, somewhat hard
- ◯ No, very hard

Next

B.D. Experiment 2: Replicator' study

Guess what another participant chose

In this study, your task is to try to *guess* what a previous Prolific participant chose in one of their decisions.

First, we will show you the instructions that the previous participant read before making their decisions. Please read them carefully as understanding them will help you with your task. The other participants' instructions will always be inside a colored box as shown below.

> These instructions are for the other participant.

Continue to read these instructions.

Next

These are the instructions that the other Prolific participant received: Page 1 of 3

> In this study, you will be making 25 decisions. In each decision, we will present you with the descriptions of 2 *charities*, and you will choose the charity you'd most prefer to make a donation to.
>
> In addition, we will randomly select one of the participants who complete this study. For this participant, we will randomly select one of their 25 decisions, and we will make a $1,000 donation to the charity that they selected in this decision.
>
> **Thus, please make your decisions thoughtfully and carefully as they could result in a substantial donation to the charity you select.**

Next

These are the instructions that the other Prolific participant received: Page 2 of 3

While we do not include the names of these charities, we include relevant information on the charity's cause and the charity's efficiency.

Here is an example charity:

Area of work: Youth Education Programs and Services
Location: Los Angeles, CA
Program Expense Ratio: 93.15%
Administrative Expenses: 15.20%
Fundraising Expenses: 7.50%
Fundraising Efficiency: $0.12
Working Capital Ratio: 0.07 years
Program Expense Growth: 12.57%
Liabilities to Assets: 1.50%

This information is collected on each charity by third-party charity evaluators. The purpose of many of these measures is to assess quality and efficiency of the charity. Below, we give you the definitions of these measures. Please read them carefully.

Area of work: Type of charity

Location: Location of the charity's operating headquarters

Program expense ratio: The fraction of the charity's total expenses (administrative expenses + fundraising expenses + program expenses) that go toward the program rather than toward administrative and fundraising expenses. This measure reflects the percent of the charity's total expenses that it spends on the programs and services it exists to deliver. A charity with a 100% program expense ratio spends *all* of its money on its charitable mission, while a charity with a 0% program expense ratio spends *none* of its money on its charitable mission.

Administrative expenses: The percent of the charity's total budget that it spends on overhead, administrative staff and associated costs, and organizational meetings. A charity with 100% administrative expenses spends *all* of its money on overhead and other administrative expenses, while a charity with 0% administrative expenses spends *none* of its money on overhead and other administrative expenses.

Fundraising expenses: The percent of the charity's total budget that it spends on fundraising, including campaign printing, publicity, mailing, etc. A charity with 100% fundraising expenses spends *all* of its money on fundraising, while a charity with 0% fundraising expenses spends *none* of its money on fundraising.

Fundraising efficiency: The amount that the charity spends in order to raise $1 in charitable contributions.

Working capital ratio: How long a charity could sustain its level of spending using its net available assets. A charity with a working capital ratio of 0 would shut down immediately without additional funds, while a charity with a working capital ratio of 20 years could sustain its current level of spending for 20 more years without additional funds.

Program expense growth: The average annual growth of program expenses. A charity with a 100% program expense growth is spending twice as much each year as it did the year before, while a charity with a -50% program expense growth is spending half as much each year as it did the year before.

Liabilities to assets: The charity's total liabilities divided by their total assets. Charities, like other organizations, need to be mindful of their liabilities in relation to their assets. This ratio is an indicator of an organization's solvency and/or long-term sustainability, with a lower ratio indicating larger long-term sustainability.

The other participant was required to correctly answer the comprehension questions below. To ensure your understanding of the instructions the other participant saw, please answer the questions below.

Imagine a charity that donates books to children in need. This charity spends $100 each day. $5 of this goes to fundraising and $65 goes to their CEO. The remaining $30 goes toward books for the children. What is this charity's program expense ratio?

Imagine a charity that just raised $365,000. This charity spends $1,000 per day. What is this charity's working capital ratio?

Imagine two charities: One charity has a fundraising efficiency of $0.20 and the other has a fundraising efficiency of $0.75. Both charities spend $100 on fundraising materials. Which charity would raise more money from this fundraising?

When you are ready, continue.

Next

After reading these instructions and correctly answering understanding questions about them, the previous participant then went on to make 25 decisions. At a random point after the fifth decision, they received the additional instructions below.
Please read this very carefully, as this directly determines your task today.

These are the instructions that the other Prolific participant received: Page 3 of 3

We have an additional task for you before proceeding.

On the next page, **we ask you to *describe* how you were selecting your preferred charities.** We call this "your description."

After this study is complete, we will recruit new participants through Prolific, and show *your description* to another Prolific participant. We will then ask this participant to *guess* the charities you selected in your past five decisions. You will receive an **additional bonus payment** if this other participant is able to match your choices. Specifically, we will randomly select one of the guesses that this other participant makes. **If this guess correctly picks the charity you chose, then you will receive a $5.00 bonus.** This other participant will also receive a bonus payment for guessing correctly.
When trying to guess the charity you picked, the other participant will see your last five decisions in random order. We will renumber the charities when they see them.

Below, please write your message to the other participant describing how you selected charities. You can click to review the instructions for writing your message and the list of charity characteristics. We have also provided a list of all the areas of work that charities could be classified into in case you want to reference them in your message.

Remember, if in a randomly picked guess the other participant guesses your decision based on your message, you will earn a bonus payment, so please write your message carefully.

Review Characteristics | Review Areas of Work

The other participant wrote a message here.

To ensure your understanding of the instructions the other participant saw, please answer the questions below.

When the other participant was writing their message, they knew that they would receive a bonus payment if:

When you are ready, click "Submit."

Submit

Only for participants in the Message condition

Your task in this study: Guess 5 choices for each of 3 participants

Your task is to try to guess previous participants' choices. We will show you the description they wrote of their decision-making process. We will show you the decisions of three participants. For each of the three participants, you will see 5 decisions. In each, you will see the two charities that the previous participant saw. Your task is to select the charity you think the previous participant chose, given their description of their decision-making process.

We will randomly select one of these decisions. If you correctly guess the chosen charity, you will earn a $5.00 bonus, and the previous participant will also earn a $5.00 bonus. If you guess incorrectly, neither of you will earn a bonus.

To ensure your understanding of your instructions, please answer the questions below.

How should you make your decisions in order to maximize your bonus payment?

Continue to start your task.

Start Task

Guess what Participant 1 chose

Participant 1 wrote this message:

I did not donate to religious charities. In general I donated based on lowest administrative expenses. If they were close, I went by fundraising efficiency.

Review Your Instructions | Review Characteristics

Participant 1: Decision 1 out of 5.

Below, please select the charity that you think Participant 1 chose in this decision given how they described their decision-making process in the message above

Area of work: Religious Activities
Location: Berrien Springs MI
Program Expense Ratio: 86.90%
Administrative Expenses: 8.90%
Fundraising Expenses: 4.00%
Fundraising Efficiency: $0.04
Working Capital Ratio: 1.11 years
Program Expense Growth: 3.96%
Liabilities to Assets: 11.80%

Choose Charity 1

Area of work: Youth Development, Shelter, and Crisis Services
Location: Greenwood Village CO
Program Expense Ratio: 79.90%
Administrative Expenses: 7.60%
Fundraising Expenses: 12.30%
Fundraising Efficiency: $0.15
Working Capital Ratio: 1.96 years
Program Expense Growth: 9.63%
Liabilities to Assets: 3.50%

Choose Charity 2

Each participant sees 5 pages like this for each of the 3 DMs, for a total of 15. Participants in the No Message condition do not see the message in the blue box.

Please answer the questions below

In general, did you find the other participants' messages comprehensible?

- ○ Yes, very comprehensible
- ○ Somewhat comprehensible (or some were and some were not)
- ○ No, I found them generally hard to comprehend

In general, how easy or hard did you find it to guess the chosen charity based on the message?

- ○ Very easy
- ○ Easy
- ○ Neither easy nor hard
- ○ Hard
- ○ Very hard

In general, did messages feel like a step-by-step process?

- ○ Yes, very much so
- ○ Somewhat
- ○ Not really

Next