Abstract

Using a laboratory experiment, we identify whether decision makers consider it a mistake to violate canonical choice axioms. To do this, we incentivize subjects to report which of several axioms they want their decisions to satisfy. Then, subjects make lottery choices which might conflict with their stated axiom preferences. We give them the opportunity to re-evaluate their decisions when lotteries conflict with desired axioms. We find that a majority of individuals want to follow the axioms and revise their lottery choices to be consistent with them. We interpret this to mean that many axiom violations in our sample were mistakes.

KEYWORDS: mistakes; rationality; normative versus descriptive theory; procedural decision making
In reversing my preference... I have corrected an error. There is, of course, an important sense in which preferences, being entirely subjective, cannot be in error; but in a different, more subtle sense they can be.

Leonard Savage (1954)

I. INTRODUCTION

An enormous experimental literature—spanning at least six decades—has shown that individuals consistently violate canonical axioms in decision theory. However, the literature has remained relatively silent on whether these violations are intentional deviations from the axioms or are simply “mistakes.” When an individual violates an axiom but would not have done so had they known they were violating the axiom, we call the violation a mistake. If violations of canonical axioms stem mainly from mistakes rather than from intentional deviations, then we can maintain confidence in the normative content of the theory, despite the fact that the theory is not descriptively accurate. However, if an individual violates the axioms because they do not want to follow them, then one should look for other “behavioral axioms” that the individual agrees with. In this paper, we provide the first incentivized experimental framework designed to detect the “subtle sense” in which individuals make mistakes in risky choice, as mentioned by Savage (1954).

Empirically identifying a mistake requires three pieces of information, reflected in the three main parts of our experiment. First, we elicit the axioms an individual wants their choices to satisfy. Eliciting preferences over axioms directly allows us to identify when an individual prefers the axiom as a principle governing all of their choices, not just in specific instances. Second, we present decision problems where the individual is likely to violate an axiom they wanted to satisfy. This part is most

1Examples include May (1954); MacCrimmon (1968); Tversky (1969); Slovic and Tversky (1974); Kahneman and Tversky (1979); Huber et al. (1982); Segal (1988); Loomes et al. (1991); Wedell (1991); Loomes et al. (1992); Camerer (1995); Birnbaum and Chavez (1997); Seidl (2002); Birnbaum and Martin (2003); Birnbaum and Schmidt (2008); Regenwetter et al. (2011); Birnbaum et al. (2016), among many others.

2This differs from the definition that takes mistakes as stochastic, following Thurstone (1927), Luce (1959), or McFadden (1973). We discuss this more in Section V.E.
similar to standard choice experiments. Finally, we observe how individuals perceive this inconsistency in their choices, and whether/how they reconcile their conflicting preferences. Since we elicit both the preference for the axiom and related lottery choices, we can present subjects with inconsistencies between their own revealed preferences, which mitigates experimenter demand effects.

We focus on six fundamental axioms in the domain of risk—-independence of irrelevant alternatives, first order stochastic dominance, transitivity, independence, branch independence, and consistency. We focus attention on the domain of lotteries, and on these axioms in particular, since many papers have shown violations of these axioms and some papers suggest that violations are a mistake while others suggest that they reflect underlying preferences. While we start in this simple domain, we emphasize that our methods could be applied to any axioms or decision-making procedures.

We find that subjects ex-ante want to follow these axioms at high rates—around 85% of subjects desire their choices to be consistent with each of the six axioms. However, as in previous experiments, subjects often violate these axioms in their lottery choices. We give subjects the opportunity to reconcile their axiom and lottery choices when inconsistent. Subjects are not required to reconcile their choices, but if they choose to do so, then they can reconcile their choices by changing their lottery decisions, by declaring they no longer want their choices to obey the axiom, or by doing a combination of these. Sample sizes vary across axioms, but aggregate results show that individuals revise about 70% of inconsistent choices, and of these revisions, about 80% change their lottery choices to be consistent with the axioms. We interpret this as subjects treating the axioms as normative and viewing their lottery choices as mistakes.

A major concern in this type of experiment is that of experimenter demand effects or other psychological concerns pushing subjects toward selecting axioms and reconciling in favor of these axioms. To control for these concerns, we include “control axioms,” which are the “opposite” of each of our axioms of interest. For example, we present subjects with the rule $c$ -transitivity (control-transitivity), which says, “If $A$ is preferred to $B$ and $B$ is preferred to $C$, then $C$ is preferred to $A$.” In comparing our main results to control axiom benchmarks, we find that subjects are much less likely to select the control axioms, doing so only about 10% of the time (compared to 85% for the axioms). Furthermore, when subjects reconcile inconsistencies in their choices,
they are much more likely to renounce the control axioms than the axioms. In doing so, they reveal that following the control axiom was the mistake, rather than their lottery choices. This gives us confidence that our observed patterns of results are not driven by experimenter demand effects or other secondary reasons. Further detail on alternative design choices are discussed in Section V.

While our results suggest that individuals do prefer to follow these fundamental axioms and that violations are often makes, we exercise modesty in generalizing our results. We do not make general conclusions that violations of the axioms in question are, definitively, mistakes. Just as it has taken decades to show where axioms are violated, it will take much more work to show where and when these violations are mistakes. Our results are suggestive of the interpretation that violations of canonical axioms can be mistakes, and we provide a framework by which to detect these mistakes. We view this paper as one step in a much larger research agenda identifying mistakes and preferences for following rules.

Moreover, it is still possible that violations of canonical axioms result from underlying preferences. For example, preferences and choices might reflect decision costs and inattention while the choices are classified as mistakes. Nonetheless, our results suggest that these choices might not reflect the axiomatic foundations given to exotic preferences. For example, individuals might make decisions as if they are motivated by a type of behavioral rule (e.g. a preference for certainty) even though they would rather satisfy a more normative rule (e.g. independence). This can only be detected by eliciting an individual's preference over rules. We view our paper as part of a literature identifying principles that an individual feels should guide their choices and when it is difficult to follow these principles. This is in a similar spirit to Oprea [2019] who studies what makes a rule complex for individuals to implement.

In line with the above, we view our paper partially as a methodological contribution and proof of concept that opens the door to a number of future research directions. Researchers can use our experimental paradigm to elicit the normative appeal of—and identify mistakes in implementing—axioms, strategies, social choice rules, and many other objects of interest. We purposefully chose simple axioms to study, but one could easily use a similar procedure to study more complicated axioms such as reduction of compound lotteries, the weak axiom of revealed preference, dynamic consistency, time stationarity, among many others. One could also compare canonical axioms to their proposed “behavioral” weakenings. In strategic games, one could
use this methodology to elicit whether individuals view obeying dominant strategies and best-responding to beliefs as normative principles for different games, even if they fail to do so successfully in practice. Researchers could also elicit attitudes toward fairness or aggregation rules in the domain of social preferences. In the case of impossibility theorems (e.g. [Arrow, 1950], these methods could be used to identify which axioms are the most desirable to relax or abandon. In this way, we view the methods here as a paradigm that can be transplanted to inform other areas of economics.

This paper contributes to a sparse literature on how individuals perceive axioms. The closest papers are MacCrimmon (1968), Slovic and Tversky (1974), and MacCrimmon and Larsson (1979). However, these studies do not incentivize the participants and do not elicit whether the individuals ex-ante want to follow the axiom in the presence of a violation. These studies also do not control for demand effects and even try to induce them in some cases. Thus, the results in our paper are the first to show, in an incentivized setting controlling for demand effects, that there is normative appeal to canonical axioms that are used to model risk preferences.

Our paper is also related to more contemporary papers that give individuals the opportunity to revise their choices (Benjamin et al., 2019; Breig and Feldman, 2019; Crosetto and Gaudeul, 2019). These papers allow subjects to change their decisions in different environments and are motivated by different research questions. However, these papers do not directly elicit subjects’ preferences over axioms. Directly eliciting axioms is crucial for our goal of identifying mistakes. Our paper also incentivizes all decisions, in contrast to some of these contemporary papers. We describe these related studies in more detail and additional literature in Section VI.

II. THEORETICAL FRAMEWORK

Before outlining the experimental design, we define the theoretical framework underlying the experiment. We presented all questions and axioms in the domain of non-negative monetary lotteries. We considered lotteries with US dollars as prizes, with potential outcomes in $X = [0,30]$. We represent the set of lotteries with prizes in $X$ by $\Delta(X)$, with strict preferences $>\text{ defined over }\Delta(X)$.$^{3}$ We denote generic prizes

---

$^{3}$Indifference and other factors such as preference for randomization are important elements of choice, and we cannot identify these in our experiment. We leave this for future work and rely on the difference between axioms and control axioms to establish our main results, since we assume that
in $X$ by $x, y, z$, and denote generic lotteries in $\Delta(X)$ by $p, q, r, s$. We represent the degenerate lottery giving $\$x$ for sure as $\delta_x$. Lastly, for a set of lotteries, $S$, we denote the set of lotteries chosen from $S$ as $C(S)$. We write $p > q$ to mean $p = C((p, q))$, or $p$ is chosen from the set of $\{p, q\}$.

Throughout the experiment, we study six fundamental axioms:

1. Independence of Irrelevant Alternatives (IIA): $p = C((p, q, r)) \Rightarrow p = C((p, q))$
   IIA states that if a lottery $p$ is chosen from the set of lotteries $p$, $q$ and $r$, then it is also chosen from the subset $p$ and $q$.

2. First Order Stochastic Dominance (FOSD): $\forall x \quad 1 - P(x) \geq 1 - Q(x) \Rightarrow p > q$
   FOSD states that if the probability of winning a prize greater than $x$ is higher in $p$ than in $q$, for all prizes, then $p$ will be chosen over $q$.

3. Transitivity (TRANS): $p > q$ and $q > r \Rightarrow p > r$
   TRANS states that if a lottery $p$ is chosen over lottery $q$, and $q$ is chosen over $r$, then $p$ will be chosen over $r$.

4. Independence (IND): $\forall \lambda \quad p > q \Rightarrow \lambda p + (1 - \lambda)q > (1 - \lambda)p + \lambda q$
   IND states that if $p$ is chosen over $q$, then the mixture of $p$ with any lottery $r$ will be chosen over the equivalent mixture of $q$ with $r$.

5. Branch Independence (BRANCH): $\lambda p + (1 - \lambda)r > \lambda q + (1 - \lambda)s \Rightarrow \lambda p + (1 - \lambda)s > \lambda q + (1 - \lambda)r$
   BRANCH states that if the mixture of $p$ and $r$ is chosen over the mixture of $q$ and $r$, then the preference will not change when $r$ is swapped out for a different lottery, $s$.

6. Consistency (CONS): $p > q \Rightarrow p > q$
   CONS states that if $p$ is chosen over $q$, then $p$ always will be chosen over $q$.

In addition to these six main axioms, we included the “opposite” of each axiom (denoted as “control axioms”). The control axioms reverse the preference relation after the implication for each of the six main axioms. We use the control axioms to test for understanding, control for potential demand effects, and disentangle explanations for selecting different axioms.

Formally, we included the following six control axioms:

---

these other considerations would be present under both types of decision rules. Additional discussion on random preferences as mistakes is in Section [V.E](#).

\[4\] Where $P(x)$ and $Q(x)$ are the cumulative distribution functions to $x$ of $p$ and $q$ respectively. For example, $P(x) = \sum_{y \leq x} p(y)$ where $p(y)$ if the probability of winning prize $y$.

\[5\] We study mixture independence rather than compound independence [Segal 1990](#). This means that $\lambda p + (1 - \lambda)r$, for example, is a reduced one-stage lottery in our lottery questions.

---

6
1. c–Independence of Irrelevant Alternatives (c–IIA): $p = C((p, q, r)) \Rightarrow q = C((p, q))$

2. c–First Order Stochastic Dominance (c–FOSD): $\forall x \quad 1 - P(x) \geq 1 - Q(x) \Rightarrow q > p$

3. c–Transitivity (c–TRANS): $p > q$ and $q > r \Rightarrow r > p$

4. c–Independence (c–IND): $\forall \lambda \quad p > q \Rightarrow \lambda q + (1 - \lambda)r \succ \lambda p + (1 - \lambda)r$

5. c–Branch Independence (c–BRANCH): $\lambda p + (1 - \lambda)r \succ \lambda q + (1 - \lambda)r \Rightarrow \lambda q + (1 - \lambda)\lambda > \lambda p + (1 - \lambda)s$

6. c–Consistency (c–CONS): $p > q \Rightarrow q > p$

We also designed six meaningless distractor rules, which were over unrelated lotteries. For example, one distractor rule is $p > q \Rightarrow r > s$ where the lotteries $p, q, r,$ and $s$ are unrelated. This rule essentially implements a random choice. We used the distractor rules only as a buffer so that subjects were less likely to notice the relationships between the axioms and control axioms. The full list of the distractor rules is in the Supplemental Appendix. When we refer to the axioms, control axioms, or distractor rules as general choice objects, we refer to them as rules, which is the language used in the experimental instructions.

We make no assumptions on preferences except for dominance in degenerate lotteries, i.e. $\delta x \succ \delta y$ if and only if $x > y$. In using the random problem selection payment mechanism, we also assume a form of monotonicity in the space of two-stage lotteries (Azrieli et al., 2019). Brown and Healy(2018) give evidence that this condition is met in a risky choice experiment similar to ours.

III. EXPERIMENTAL DESIGN

Identifying a mistake requires three pieces of information: eliciting the axioms an individual wants their choices to satisfy, identifying violations of these axioms, and studying how discrepancies in these preferences are reconciled. Our experiment consisted of three main blocks to elicit these three pieces of information. First, we overview these blocks and discuss the underlying design choices in each block. We present more details for each block in the subsequent subsections.

---

6 This assumption allows us to represent multiple switches on a price list as a violation of transitivity. It is also relevant for FOSD violations.

7 We included an additional part designed to elicit ranking over axioms and willingness to pay for the opportunity to reconcile choices. We defer explanation of this to Appendix Section IV.A.
III.A. Overview

We summarize the most important design choices and brief reasoning below. We discuss our design choices in light of forgone alternative methods in Section \( \text{V} \)

1. All decisions, including the choice to follow an axiom, are incentivized.

2. We directly elicit an individual’s preference for satisfying a decision rule.

3. Control axioms capture demand effects, confusion, and other latent tendencies to follow rules while having the same form for all axioms.

4. The opportunity to reconcile choices is neutral so that one can make changes to axiom choice, lottery choice, both, or have choices remain inconsistent.

In Block 1, we elicit an individual’s preference for satisfying decision rules. Eliciting preferences over rules presents many challenges, such as ensuring that subjects understand the decision rules, incentivizing subjects’ responses, and controlling for demand effects. To help subjects understand the axioms, we explained them using simple colored circles rather than using their mathematical expressions. To incentivize selection of an axiom, we presented them as “decision rules” making a relevant choice on a subject’s behalf. For example, a subject who wants their choices to be transitive, and who chooses \( A \) over \( B \) and \( B \) over \( C \), would have the choice of \( A \) over \( C \) automatically made for them if the transitivity axiom were chosen for payment. Finally, to control for experimenter demand effects and other motivations for selecting rules, we include the “opposite” of each axiom, which we refer to as our “control axioms.” Differences between selection rates of the axioms and control axioms cannot be explained merely by experimenter demand effects, subjects not wanting to make choices on their own, responsibility aversion, etc. since the control axioms are presented and incentivized in the same manner as the main axioms.

After eliciting axiom preferences, in Block 2 we present lottery choices which offer the possibility of individuals violating an axiom they wanted to satisfy. Finally, in Block 3, we observe how individuals perceive inconsistencies in their choices, and whether/how they reconcile their conflicting preferences. To mitigate experimenter demand effects, we provide subjects with a neutral reconciliation opportunity. That

\[ \text{footnote}{8}{If this subject did not select the transitivity axiom, they would make the choice between } A \text{ and } C \text{ on their own.} \]
is, subjects could make their choices internally consistent by renouncing the axiom, by changing their lottery choices to be consistent with the axiom, or some combination of these things. There is no default direction for this reconciliation opportunity, and we also allow subjects to keep their choices inconsistent. This not only allows us to identify a mistake, but we can see, from the subject’s own perspective, whether the mistake was in the axiom choice or in the lottery choice. We also allow subjects to revise inconsistencies with any control axioms they selected. Differences in reconciliation patterns between the axioms and control axioms cannot be explained by experimenter demand effects or any other spurious considerations since the axiom and control axiom reconciliations are presented in the same manner.

We describe the choices and payment mechanisms in detail below. To overview the payments, subjects could be paid for one of four possibilities: original rule choices (Block 1), original lottery choices (Block 2), revised rule choices (Block 3), or revised lottery choices (Block 3). The incentivization procedures are the same for original and revised rules, and are the same for original and revised lotteries. Rules are incentivized by applying them on a set of lotteries and paying subjects what the rule prescribes selecting. If individuals do not want to follow a rule, instead they make the lottery choices themselves. Original and revised lottery choices are incentivized in the standard manner, by paying subjects a realization from the lottery they selected. All payment uncertainty was resolved using physical randomization devices, in particular two ten-sided dice. The choice of which question would be paid is based on random chance. Subjects were paid at the end of the experiment, regardless of which decision was selected for payment.

III.B. Block 1: Rule Choices

The objective in Block 1 was to elicit whether subjects want their decisions to satisfy canonical choice axioms. The first challenge is incentivizing the rule choice so that subjects select all of the rules they view as desirable and do not select any others. We did this by presenting rules as logic statements and asking subjects to decide whether they prefer the rule to make choices for them or whether they would rather make the relevant choices themselves. If the subject preferred a rule to make decisions for them and the rule was selected as the payoff-relevant decision, then we applied

---

9Subjects were not allowed to choose between these two options until at least 30 seconds had passed. This design feature encourages subjects to consider the rules carefully before deciding.
the rule to a set of lotteries where the it has implications. The subject was paid a realization of the lottery prescribed by the implications of the rule. If the subject did not select a rule to make decisions for them and the rule was selected as the payoff-relevant decision, then they would make the relevant choice on their own.

For example, if IIA were chosen for payment, then we would present the subject with three lotteries corresponding to $p$, $q$, and $r$. The subject would choose their most preferred lottery from the set \( \{p, q, r\} \). Suppose the chosen lottery is $p$. The subject would be paid from the binary decision problem involving the chosen lottery and some other lottery, e.g. \( \{p, q\} \). If the subject wanted to follow IIA, then we would automatically implement the choice of $p$ over $q$ for them, as prescribed by IIA, and would pay them a realization of the lottery $p$. If the subject did not want their choices to satisfy IIA, then we would present them with the choice of this lottery, $p$, versus one of the other lotteries, e.g. $q$. We would pay them a realization from whichever lottery they chose in this second decision.\footnote{During payment, subjects were not told which rule was being implemented before they made the input choices. If we had told them which rule was being implemented, they could answer the input questions “opposite” their true preferences for the control axioms and still receive their truly-preferred alternative.}

![Figure I: Rule representation of IIA](image)

**Notes:** We represent axioms in this manner. Colored circles represent any possible lotteries with payoffs from $0–$30.\footnote{We also included a written description of the rule on the subjects’ screens under the abstract depiction. We described the rule using fruits which match the colored circles. For example, we say “If you choose a blueberry over an apple and a grape, then we would choose a blueberry over an apple for you.” Subjects are reminded, though, that the circles represent money and not consumption goods, and there are no complementarities or tastes involved in mixtures.} Subjects choose whether to have this rule make choices for them or instead make choices on their own.

The second challenge in eliciting preferences over rules lies in making the domain of the rules accessible and easy for subjects to understand, while still retaining the broad implications on choices. We chose to present the decision rules in a simple way,
representing lotteries by colored circles. Subjects were told that the colored circles represent monetary lotteries but they did not know the exact lotteries associated with each rule. We inform subjects that the lotteries could have payoffs from $0–$30, with any probabilities from 0%–100%. Again, we use IIA as an example show how we present the rules to subjects in Figure I. In Appendix Section E, we show how we represent the other five axioms in rule format. Combining the axioms, control axioms, and distractor rules, subjects made eighteen decisions in Block 1.

We incentivized the rules this way to encourage subjects to select exactly the axioms which always should be true on the relevant domain of lotteries. Thus, subjects should select the rules that they want to describe all of their choices, regardless of the specific lotteries and payments. A subject who does select a rule reveals that they want to make decisions according to the rule, since it can be applied over any lotteries in the domain. However, the interpretation is less clear for subjects who do not select a rule. A subject who agrees with a rule but believes their choices will align with the rule anyway has no strict incentive to select the rule, aside from the time and effort cost of making choices on their own.

In Appendix Section B, we present results from another treatment where subjects had to pay a small cost, $1, to make the choice on their own. This makes it strictly costly not to select a rule, eliminating this concern. We find that the rules are selected slightly more often in this treatment, responding to the incentives, but all qualitative results remain unchanged. This indicates that the vast majority of subjects who do not choose a rule are not doing so out of indifference, but rather because they do not find it desirable.

III.C. Block 2: Lottery Choices

Given that our main interest is in studying how individuals reconcile inconsistent choices, we selected lottery questions from previous papers which found violations of the axioms. We do not focus on the specifics of the lotteries, but we picked questions to maximize axiom violations. Our intention is not to compare violations across axioms since the questions are not representative of the axiom more generally. The full set

\[\text{III.C. Block 2: Lottery Choices}\]

Given that our main interest is in studying how individuals reconcile inconsistent choices, we selected lottery questions from previous papers which found violations of the axioms. We do not focus on the specifics of the lotteries, but we picked questions to maximize axiom violations. Our intention is not to compare violations across axioms since the questions are not representative of the axiom more generally. The full set

---

12 Here, however, the interpretation is less clear for subjects who do select a rule. A subject who selects a rule does not necessarily indicate they always want to follow the rule. It could be that they want to follow the rule “most” of the time, so in expectation, they believe it is not worth paying $1 to make choices on their own.
of questions and descriptions can be found in Appendix C.

We displayed the lotteries simply by reporting the probabilities and payoffs of each possible outcome, as shown in Figure II. Subjects saw the lotteries on their screens as below and made their choices by selecting the rectangular button corresponding to their preferred option. Altogether, subjects make choices from thirty-three binary or trinary decision problems in Block 2.

One issue when studying preferences over axioms is that if an individual views many axioms as normative, then the axioms combined may have unexpected implications across questions. For example, continuity, completeness, transitivity, and independence together imply that there is an expected utility representation consistent across all questions. To ameliorate this concern, we chose lotteries so that the questions have no direct implications across multiple axioms. In particular, we did not use any of the same lotteries to target more than one axiom. The lotteries used in FOSD1, for example, are completely distinct from lotteries used in any other question. Therefore, there are no direct implications of the axioms across questions, and each decision applies only to its targeted axiom.

![Option A: 50% chance of $3 50% chance of $15  Option B: 25% chance of $5 75% chance of $12](image)

**Figure II:** Representation of lotteries

**III.D. Block 3: Reconciliation**

After subjects ranked all of the axioms, we presented them with every inconsistency between their lottery choices and selected axioms or control axioms. For example, a subject who selected IIA in Block 1 but violated IIA with their lottery choices in Block 2 saw these choices side by side on their screen. We will focus on the difference in reconciliation patterns between the axioms and control axioms as our main result. Given that we use the control axioms as our benchmark, and due to concerns about experiment duration and cognitive load, we do not include “placebo” revision oppor-
tunities where subjects could revisit choices that were already internally consistent. We discuss this more in Section V.B.

On subjects’ screens, we highlighted the original choices that the subject made for rules and lotteries. We match the subject’s lottery choices to the colors of the rule when presenting the reconciliation opportunity so the subject could better understand how the rule mapped onto their choices. We also include a written explanation of why choices violated the rule and how they could be reconciled. We used neutral language in describing the violations. We phrase any inconsistency in rule and lottery choice by saying that the rule would have chosen something different for the subject than what the subject chose for themselves. We provide a screenshot in the Appendix A and reproduce an example below in Figure III. The language used to describe violations with the control axioms is identical to the language used to describe violations with the main axioms. We also provide a control axiom reconciliation screenshot in Appendix A.

![Figure III: Example of Reconciliation Screen](image)

**Notes:** The options highlighted in grey indicate subjects’ original choices in Blocks 1 and 2. For example, this subject selected IIA in Block 1, but chose “black over grey and white” in one question and chose “grey over black” in another question. Below this, subjects saw an explanation of why the rule would have selected something different than what they chose for themselves.

Subjects could change *any* of their choices, or could leave them as they were. We impressed upon subjects that they could change any of their lottery choices, could unselect the rule, could do both, or could leave choices inconsistent. For example, suppose, as in Figure III an individual selected IIA as a decision rule and then chose lottery \( p \) from \( \{p, q, r\} \) and \( q \) from \( \{p, q\} \). The individual could unselect the rule, could change their selection from \( \{p, q, r\} \), could change their selection from \( \{p, q\} \), could do combinations of these, or could do nothing. As a result, there was no default direction
for any potential experimenter demand effect, which is a key feature in our design. If subjects were making their choices internally consistent simply due to experimenter demand effects, then we would expect that they would be equally likely to unselect the axiom as they would be to change their lottery choices.

Our key assumption is that when an individual revises their axiom and lottery choices to be consistent, this reveals that the original choices were a mistake. We believe this is a reasonable assumption since the individual has strictly more information about the implications of their previous decisions. However, as mentioned earlier, revisions could also be caused by demand effect. For this reason, it is critical that there is some benchmark behavior to compare reconciliation behavior. Thus, for our analysis we focus on the difference in revision patterns between the axioms and control axioms.

While the reconciliation opportunity occurs on a single screen, any choice on the screen has an independent chance of being selected for payment. For example, consider the reconciliation opportunity for IIA. A subject could be paid for their revised rule choice, their revised choice from \(\{p, q, r\}\), or their revised choice from \(\{p, q\}\). Each choice on the screen is paid in the same manner as the original choices were paid in Blocks 1 and 2. Furthermore, the subject’s original choices from Block 1 and Block 2 were not overturned by this reconciliation opportunity and still could have been chosen for payment.\(^{13}\)

Subjects had the opportunity to reconcile choices inconsistent with each of the six axioms and the six control axioms.\(^{14}\) Subjects reconciled each violation independently. That is, a subject who selected IIA and violated it on two separate occasions had two separate opportunities to reconcile the violations, rather than reconciling all choices together.\(^{15}\) We did this to encourage subjects to analyze each choice in isolation, and to reduce cognitive demand in the reconciliation stage.

\(^{13}\)Choices which were not violating a rule also could be paid again as reconciliation choices to maintain equal probability of all rules and lotteries being paid. In this case, we paid the subject based on their original rule or lottery choice.

\(^{14}\)We did not have subjects reconcile \(c\text{--TRANS}\) with the price list, as we could not explain how to make the price list completely intransitive. We did not have subjects reconcile the meaningless distractor rules given that there is no natural way to present the violating choices.

\(^{15}\)This also means that the reconciliation was not dynamic. That is, a subject who selected both IIA and \(c\text{--IIA}\) in Block 1, and then violated IIA in Block 2, may have reconciled these choices to be consistent with IIA. In doing so, this might lead them to be inconsistent with \(c\text{--IIA}\). We did not present them this subsequent reconciliation. The reconciliation opportunities were fixed and determined by their choices in Blocks 1 and 2.
Subjects also had the opportunity to reconcile inconsistencies when they chose both the axiom and control axiom. For example, a subject who chose both IIA and c−IIA in Block 1 would also see these rules side by side on their screen and choose which, if any, to keep selected. The subjects were not presented with their lotteries during this reconciliation opportunity. Again the language here was neutral and simply said that these two rules make opposite choices. These decisions were incentivized in the same way as other revised rule choices.

The number of reconciliation opportunities varied per subject, based on number of violations and on number of axioms and control axioms selected. On average, subjects had 5.5 reconciliation opportunities. The number of the reconciliations ranged from zero to twenty-two.

We analyze data from 110 subjects, primarily undergraduate students at the Ohio State University where the sessions took place. We programmed the experiment using z-Tree (Fischbacher, 2007), and recruited subjects using ORSEE (Greiner, 2015). Sessions lasted about one hour, and subjects earned $15 on average, including a $7 show-up payment. Instructions are included in a Supplemental Appendix.

IV. Results

Figure IV shows the percentage of subjects who selected each axiom in Block 1, broken down by whether a subject selected the axiom only, the axiom and the control axiom, only the control axiom, or neither. In aggregate, FOSD is most popular, selected by 90% of subjects. For the remaining axioms, 85% select Consistency, 83% select Transitivity, 83% select IIA, and 83% select Independence, and 82% select Branch Independence. More importantly, the axioms are much more often selected than the control axioms. In particular, 15% selected c−Branch Independence, 14% selected c−Independence, 12% selected c−IIA, 11% selected c−Transitivity, 10% selected c−FOSD, and 5% selected c−Consistency.

Interestingly, FOSD is the most popular axiom while c−FOSD is among the least popular control axioms. Similarly, Branch Independence is the least popular axiom while c−Branch Independence is the most popular control axiom. This might indicate that there are some patterns to how subjects perceive the axioms. FOSD is most “obviously” desirable, and therefore c−FOSD is obviously not desirable. The opposite is true for Branch Independence. This suggests some features of axioms might
be more or less compelling to individuals. It is also interesting that both FOSD and Branch Independence involve “mixing,” so it is not that case that individuals are just averse to, or confused by, mixing. From our data, however, we cannot say whether this is driven by differences in understanding of the axioms or differences in preferences over the axioms.

Our aggregate results are reflected in the individual-level rule selection rates. 60% of subjects selected all six axioms, and 65% of subjects never selected a control axiom. Among individuals who ever select a control axiom, it is most common for individuals to select only one. Therefore, we have confidence that subjects understand the decision rules and incentivization procedure. More individual-level results are provided in the Appendix. 

Overall, we conclude that individuals do view these axioms as desirable methods for choosing lotteries since they are chosen at much higher rates than the control axioms.

---

Figure IV: Percentage of Subjects Selecting Each Rule in Block 1

---

\[^{16}\] In particular, the number of axioms and control axioms selected on an individual level is reported in Table IX. In addition, we report correlations of selecting the various axioms. We find that all pairwise correlations are significantly positive, but we do not detect any obvious patterns.
ions. While these same individuals will go on to violate the axioms in their choices, they express a preference for their choices to satisfy these axioms as general principles.

**Result 1.** Nearly all individuals reveal a preference for their choices to satisfy canonical choice axioms. These axioms are selected at higher rates (≈ 85%) than their “opposites” (≈ 10%).

Given that subjects prefer to satisfy these axioms, the question remains as to how individuals respond when their choices do not. Among those who select the respective axiom, 85% of subjects violated FOSD, 75% violated Independence, 46% violated Consistency, 43% violated transitivity, 38% violated IIA, and 24% violated Branch Independence. Over 85% of subjects who violated an axiom indicated in Block 1 that they wanted to follow the axiom, so “wanting” to follow a rule does not ensure that a subject can or will follow the rule. Our main focus is on the difference in reconciliation frequencies and patterns between the axioms and control axioms.

Figure V shows choices in the reconciliation block (Block 3), broken down by question. While we do not focus on specific questions since they vary on many dimensions, there is observable heterogeneity across questions and across axioms. A non-trivial percentage of inconsistencies are not reconciled, which we discuss below. However, of those who do change their choices, it is far more common for individuals to change their lottery choices to be consistent with the axiom than to unselect the axiom.\(^{17}\) Across all questions, 79% of revisions are making lottery choices consistent with the axiom, rather than the opposite. Naturally, this is significantly different from a 50–50 split in the direction of revisions \((p = 0.0000)\).

Table I breaks this down by axiom. In column two, we report the percentage of individuals who changed any of their choices in the reconciliation. In columns three and four, we report the conditional percentages of revising the lottery to follow the axiom versus unselecting the axiom, excluding those who change both previous decisions. Note, the sample sizes vary widely as individuals violated some axioms more than others, and some axioms had more questions than others.\(^{18}\)

\(^{17}\) In our main treatment, there are no instances of subjects changing their lottery choices, while keeping the axiom selected, in such a way that the revised lottery choices are inconsistent with the axiom.

\(^{18}\) For example, there were four FOSD questions, and 85% of subjects violated FOSD at least once. On the other hand, there was only one Branch Independence question and 24% of subjects violated the axiom.
Figure V: Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting Axiom

Notes: Here, “change lotteries” means the individual changed their lottery choices to match the rule, “unselect rule” means they no longer want to follow the rule, “change both” means both lottery and rule choice changed, and “keep inconsistent” means they did not revise anything. There were no subjects who revised lotteries and rules in a way to be inconsistent.

<table>
<thead>
<tr>
<th>Axiom</th>
<th>% Revised Either</th>
<th>Conditional % Revise Lottery</th>
<th>Conditional % Unselect axiom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any (n=782)</td>
<td>63%</td>
<td>79%</td>
<td>21%</td>
</tr>
<tr>
<td>IIA (n=116)</td>
<td>76%</td>
<td>99%</td>
<td>1%</td>
</tr>
<tr>
<td>FOSD (n=359)</td>
<td>44%</td>
<td>66%</td>
<td>34%</td>
</tr>
<tr>
<td>TRANS (n=46)</td>
<td>80%</td>
<td>94%</td>
<td>6%</td>
</tr>
<tr>
<td>IND (n=159)</td>
<td>50%</td>
<td>78%</td>
<td>22%</td>
</tr>
<tr>
<td>BRANCH (n=22)</td>
<td>59%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>CONS (n=80)</td>
<td>91%</td>
<td>100%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table I: Percentage of Violations Revised and Direction of Reconciliation

Notes: The second column gives the percentage of violations that were reconciled. The third column reports the percentage of these reconciliations where subjects revised their lottery choices to be consistent with the axiom, and the last column reports the percentage of these reconciliations where subjects revised their rule choices. This excludes the minority of subjects who changed both of their axiom and lottery choices.
We observe heterogeneity in the tendency to revise inconsistencies. It is interesting to note that FOSD, IND, and BRANCH are the least likely to be revised, and these are the three axioms which involve “mixing.” From our data, we cannot say whether this reflects subjects’ preferences related to these axioms, or whether it is simply harder for subjects to understand why their choices violate axioms that involve mixtures. This is especially interesting since FOSD is the most frequently selected axiom in Block 1. This shows that even though individuals may want to follow an axiom, this may not translate to them making choices consistent with it even when given an explanation of how the axiom applies to a decision problem. We leave further investigation of reconciliation properties for specific axioms to future research.

We use the reconciliation of the control axioms as a benchmark for the direction of reconciliation. We might worry that individuals change their choices to match the rules as a form of experimenter demand effect. We tried to eliminate this in the design—since subjects could make their choices consistent either by changing their lottery choices or by un-selecting the rule, there is no obvious direction for the demand effect. If choices were made internally consistent solely due to experimenter demand, we would expect it to be equally likely for subjects to change their axiom and lottery choices. This is clearly not the case for our axioms, as subjects are more likely to revise their lottery choices to be consistent with the axiom. However, we show that revisions for the control axioms closely follow a 50-50 split between changing lotteries and unselecting the rule.

Figure VI shows the same breakdown of revised choices conditional on subjects selecting the control axioms. Across all questions, 44% of revisions change lottery choices to be consistent with the control axioms, not significantly different from a 50–50 split (p = 0.311). This is in significant contrast to the 79% of instances where subjects reconcile in favor of the main axioms (Wilcoxon Ranksum p = 0.0000). The 44% of revisions that change lottery choices to be consistent with the control axioms captures any latent tendency to follow rules. For example, subjects might think that the experimenter wants them to follow rules, or subjects may prefer that rules make choices for them rather than making choices on their own. Nevertheless, the difference between the control axioms and our main axioms suggests that a significant portion of our revisions cannot be explained by these other factors.

---

19 We thank Yoram Halevy for this observation.
Figure VI: Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting Opposite Axiom

Notes: Here, “change lotteries” means the individual changed their lottery choices to match the rule, “unselect rule” means they no longer want to follow the rule, “change both” means both lottery and rule choice changed, and “keep inconsistent” means they did not revise anything. There were no subjects who revised lotteries and rules in a way to be inconsistent.

One might still worry, however, that individuals who select the control axioms are systematically different from those who select the axioms. For instance, they might not understand the rule representations or payment incentives, which could explain why they selected the control axiom in the first place. We have no reason to believe these factors would lead to systematic differences in reconciliation patterns, but we can rule this out further by looking at individuals who select both an axiom and corresponding control axiom. Conducting the same analysis as above, restricted to this subsample, we find the results unchanged. When reconciling violations of the axioms, individuals revise their choices 69% of the time, and 65% of revisions change lottery choices to be consistent with the axiom. When reconciling violations of the control axioms, these individuals revise their choices 72% of the time, and only 36% of revisions follow the control axiom, which is significantly lower than for the main
axioms ($p = 0.0149$).

Furthermore, Figure VII shows the rule reconciliation pattern for these individuals. That is, we look to see how individuals reconcile their rule choices when they selected both the axiom and corresponding control axiom. They had the opportunity to unselect the axiom, unselect the control axiom, both, or neither. Within that sample, we find individuals still favor the main axioms. Among individuals who un-select one of the rules, over 89% of them un-select the control axiom. That is, when individuals are faced with two decision rules that prescribe opposite choices, they realize this and abandon the less-sensible rule.

**Figure VII:** Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting axiom and control axiom

Notes: Here, “unselect c-axiom” means the individual kept the axiom selected but unselected the control axiom, “unselect axiom” means they kept the control axiom selected but unselected the axiom, “keep both” means they kept both the axiom and control axioms elected, and “unselect both” means they unselected both the axiom and control axiom.

We conclude that a majority individuals who wanted to follow the axioms but violated them made a mistake in their lottery choices. There is a baseline tendency to keep choices inconsistent, as we will discuss below. However, among reconciliations,
the axioms are usually followed.

**Result 2.** *Individuals violating canonical axioms often change their choices to be consistent with the axiom* (≈ 79% of revisions). *Individuals violating control axioms are less likely to do so* (≈ 44% of revisions).

Finally, Figures V, VII, and VI show that about 30% of subjects keep their choices inconsistent across these revision opportunities. Inconsistencies with the axioms are revised 63% of the time and inconsistencies with the control axioms are reconciled 67% of the time (\(p = 0.37\))\(^{20}\)

While this might seem odd at first blush, there are a few reasons why individuals might keep their choices inconsistent. The most obvious to us is simple effort cost. Subjects have already thought about these decisions and chosen what they prefer. Revising choices is costly in terms of time and cognitive effort, and individuals may view the cost as too high. To test this hypothesis, we look at the first and last revision opportunity that subjects faced. Averaged across all subjects, we find that choices are left inconsistent 34% of the time in the first reconciliation opportunities, while they are left inconsistent 46% of the time in the last opportunity (Fisher exact \(p = 0.052\))\(^{21}\) These revision opportunities vary on other dimensions, so more research is required to test this hypothesis directly. However, evidence suggests that keeping choices inconsistent is driven, in part, by attention and choice fatigue, rather than by a calculated desire to hold both preferences despite being internally inconsistent.

**Result 3.** *Individuals keep their choices inconsistent in about one third of all reconciliation opportunities. We find suggestive evidence that choice fatigue contributes to subjects’ willingness to maintain inconsistencies in their choices.*

While we could look at detailed comparisons in original and revised lottery choices, for example whether revised choices become more or less risk averse, our experiment is not designed for these questions. We chose the lottery questions in order to max-

---

\(^{20}\)This is true even among individuals who have a strict preference for re-evaluating their decisions, as measured by the willingness to pay elicitation we describe in Appendix Section IV.A.

\(^{21}\)This is even stronger in our $1 cost treatment, where first revisions are left inconsistent 34% of the time and last revisions are left inconsistent 64% of the time (\(p = 0.000\)). Individuals had more revision opportunities in this treatment, on average, since they select axioms more often due to the $1 cost of not selecting the axiom. The fact that choice fatigue seems to increase in this treatment where there are more revision opportunities strengthens the hypothesis that inconsistencies are due to these related factors.
imize violations of the axioms, and therefore the questions are in no sense representative of the violations and revisions we might see more generally. However, we believe our methodology could be very useful in answering these types of questions in future research. For a step in this direction, see Benjamin et al. (2019) who study risky investment decisions before and after reconciliation opportunities.

V. DISCUSSION OF ALTERNATIVE DESIGN CHOICES

We carefully designed our experiment to allow for clear interpretation of mistakes with minimal complexity for subjects. We briefly discuss alternative design choices that we could have made and our justifications for not doing so. We believe this exercise will be particularly useful for researchers who wish to transport our framework into other choice domains.

V.A. Eliciting Rule Preference

We chose to elicit subjects’ preferences over rules directly. Other papers in the literature have allowed for reconciliation opportunities, but without eliciting preferences over rules. These experiments either explain to subjects that their choices violate a given rule or repeat decision problems with or without presenting the original choice. We felt that our design would have a clearer interpretation and would be less likely to generate experimenter demand effects relative to these other alternatives for the following reasons.

If a given decision rule is only explained by the experimenter, then this never gives the subject a chance to voice approval or disapproval of the decision rule. We felt this would be more likely to lead to subjects changing their choices out of “embarrassment” since the subject is explained the rule by an authority on decision making (e.g. the experimenter). In contrast, eliciting a preference for the decision rule from the subject allows us to know whether the subject prefers the axiom always. Furthermore, it allows the subject to effectively “give themselves advice” when we present them with their lottery choices related to the rule. Thus, rather than explaining rules to subjects, subjects themselves select the rules they want to follow and are able to reevaluate these decisions in light of conflicting evidence in their preferences.

If the experimenter offers the individual a repeat decision without presenting the original choice, then one must be concerned about which choice is the “revision.” To
better understand this, since the subject does not know the decision rule associated with a given question, the researcher cannot say whether the initial choice or revised choice is the one more favored by the individual; we can only see that the choices are different. In contrast, because we elicit the individual’s preference for decision rules, we can more clearly interpret the later choices as the revised choices since the axioms give subjects more information about their own preferences. Eliciting axioms also allows us to know what rules (ex-ante) a subject wants to follow, as well as knowing their preference for decision rules “ex-post,” after confronting compelling counterexamples to the rule.

Finally, eliciting preferences over decision rules gives us a richer dataset on subjects’ preferences. For example, by selecting a rule, a subject reveals that they prefer all of their choices to be consistent with the decision rule on the relevant domain. Without eliciting the axiom preference, we cannot make a claim about an “overall” preference for following the axiom. For any of the alternative schemes above, even when a subject reconciles inconsistent lottery choices to be consistent with a rule, we could only interpret this as wanting to follow the rule for those particular questions. In contrast, our design allows us to elicit global information about preferences for a given domain.

V.B. Reconciliation Opportunities

We carefully designed our reconciliation opportunity to be neutral for subjects. In particular, there was no default direction to reconciliation; we presented a subject simultaneously with both their rule choice and lottery choice to reduce experimenter demand effects. A subject could unselect a rule, change their lottery choices, a combination of these changes, or they could doing nothing. A drawback of this is that allowing a subject to change potentially all previous decisions likely places a higher cognitive demand on our subjects. However, we mitigate this through clear explanation of the relationship between rules and lotteries during the reconciliation opportunity.

We did not include “placebo” reconciliation opportunities. A “placebo” reconciliation opportunity would allow a subject to change choices that were already consistent with a rule. This design choice was made since we expected individuals would experience choice fatigue from facing a large number of reconciliation opportunities. In-
Indeed, we find evidence that individuals revise their choices less in later reconciliation opportunities. Thus, including placebo reconciliation opportunities would have only increased the cognitive load on subjects and reduced our ability to detect mistakes. Furthermore, we did not allow individuals to select a rule that had not been chosen originally, even if they satisfied the rule in their lottery decisions. Since we interpret selecting an axiom as a global preference for satisfying it, seeing a single set of choices that are consistent with the axiom should not affect an individual’s global preference for satisfying the axiom elsewhere. This remains to be tested empirically.

Finally, we had subjects reconcile each question that violates a rule independently, rather than doing “batch” reconciliations for a given rule. This allows for subjects to make exceptions to the rule based on “what is more rational to do in this instance,” as discussed by [Gilboa et al. (2009)](#). Moreover, we felt that allowing batch reconciliations while maintaining neutrality of the reconciliation opportunities would place more cognitive demands on the subjects.

**V.C. Control Axioms**

We use the control axioms as a benchmark to the elicited preference over axioms. Recall, the control axioms are an “opposite” of the axioms of interest. These control for experimenter demand effects since any argument that the axioms are chosen because they come from an authority also applies to the control axioms. We chose the control axioms to be the opposite of our main axioms for two reasons. First, this provides a standardized form for the benchmark across all six axioms. The control for each axiom is its opposite, rather than choosing different types of controls for different axioms. Second, in doing so, the control axioms control for any “axiom-specific” confusion or other bias. For example, if one believes that our visual representation of IND is driving subjects’ preferences for following it, then this would also be true for the control axiom, c-IND, since it is displayed in a similar form.

One could have imagined many other benchmarks. For example, we could have used relaxations of the axioms, heuristics which are “mostly” true, etc. We chose not to do these alternatives for a few reasons. In particular, we wanted to start simple. The control axioms are arguably the “most obvious” benchmark since they are the opposite of our main axioms and take the same form regardless of the axiom. Relaxations of the axioms typically require more “conditioning” and domain restrictions.
which could complicate the presentation of the axioms and subjects’ understanding of the representations. Moreover, relaxations or heuristics that are “mostly true” do not take the same form for different axioms and thus would not serve as a consistent benchmark. We believe these alternative implementations introduce interesting complications for future research to investigate.

V.D. Incentivization Scheme

We believe it is important to incentivize decisions, but the method used to incentivize preference for decision rules is non-trivial. We chose to elicit an individual’s preferences to follow the decision rule relative to making a decision on their own. This avoids interpretation issues that would arise with different incentive schemes, such as implementing a random choice or the opposite choice if subjects prefer not to follow an axiom. It also allows us to identify whether an individual wishes to satisfy an axiom all of the time relative to a common benchmark (e.g. their decision).

There is no cost to decide on one’s own in the main experiment, so that subjects who choose to follow an axiom strictly reveal that they believe the rule should always be true. Our robustness treatment introduces a $1 cost to making the decision on one’s own, and our main results are stronger, if anything. We do not have a treatment where it is costly to follow a decision rule. If an individual chooses to follow a decision rule when it is costly to do so, this would tell us something about the individual’s beliefs about their ability to follow the rule in their own decisions, and/or something about their beliefs about the questions they will face. The interpretation and implementation of the reconciliation opportunity would differ in this case, since subjects might want to know when they do satisfy an axiom in their choices, leading them to unselect it as a costly decision rule. Thus, this would give interesting information, but it is outside the scope of this paper.

V.E. Random Mistakes

An alternative perspective on mistakes often uses the framework of additive random utility models (Thurstone 1927; Luce 1959; McFadden 1973). However, this perspective is more subtle than it first appears. For example, any mistake generated from additive random utility models can be described by a non-expected utility model of risk (Hofbauer and Sandholm 2002; Allen and Rehbeck 2019). Thus, from
a modelling perspective, one cannot differentiate between random behavior being a “mistake” or a preference to randomize. In contrast, by eliciting a preference for a decision rule, we at least recover some ex-ante indicator of whether an individual prefers the decision rule.

V.F. Choice and Number of Lotteries

We chose to use a few questions per axiom, based on classic violations in the literature. It would be interesting to do more exhaustive analysis on each axiom to get an overview on where mistakes occur most often, but we leave this for future work.

VI. Literature Review

Naturally, our paper relates directly to the large literature on violations of decision theory axioms. Many papers have found violations of the axioms we study. We confirm these violations, but our experiment suggests that some of these violations might be better thought of as mistakes. Individuals do want their choices to satisfy these axioms, and, when given the opportunity for reflection and revision, individuals often prefer to change their choices to be consistent with the axiom.

We are not aware of any papers that incentive-compatibly elicit the axioms an individual wants to follow. Even without incentives, we are not aware of any papers that elicit axioms and allow for reconciliation with violating choices, which is necessary to detect whether an individual made a mistake. There are three related papers in economics which come closest to our motivation: MacCrimmon (1968), Slovic and Tversky (1974), and MacCrimmon and Larsson (1979). Since our work builds off these pioneering studies, we describe them in detail.

MacCrimmon (1968) asked questions to thirty-eight business executives designed to induce potential violations of five postulates of decision theory. After making decisions, the postulates of choice were discussed with the executives and the executives were allowed to change their choices during the discussion. Within the experiment, most executives were either consistent with the postulates of decision theory or changed their choices in favor of the postulates after some discussion, consistent with our results. While an ambitious study, the unstructured nature of the discussion, lack of incentives, and specialized subject pool are limitations to external validity of the results. Furthermore, the postulates were presented to subjects by the
experimenter in the reconciliation opportunity. A main contribution of our experiment is that subjects themselves choose both the axioms and the lotteries, so the reconciliation is less prone to experimenter demand effects and other outside motivations.

In a related study, Slovic and Tversky (1974) ask unincentivized lottery questions related to the Allais and Ellsberg paradoxes to undergraduate students. After making their decisions, subjects were presented with “advice” in the form of explained arguments for and against the independence axiom and the sure thing principle. Subjects were told that the arguments came from prominent decision theorists. After reading the arguments, subjects were asked to reconsider their original lottery choices. Overall, they find few subjects changing their decisions to be consistent with the advice. In addition to being incentivized, a contribution of our design is to elicit individuals’ preferences over these axioms, rather than presenting them as advice or arguments from others. This allows individuals to see the inconsistencies within their own preferences, rather than deciding whether to follow the logic from an outsider. This could be one explanation for why our subjects are more likely to revise their choices compared to Slovic and Tversky (1974).

MacCrimmon and Larsson (1979) asked graduate students to indicate their level of agreement with 20 decision rules related to the axioms of expected utility theory. Additionally, they had subjects make unincentivized lottery choices and indicate any decision procedures involved in making the choices. They find that individuals view the decision rules favorably but often violate them. In this experiment, there is no reconciliation stage. We design our experiment to carefully incentivize these choices and to study how individuals reconcile discrepancies between their choice of lotteries and decision rules.

A number of contemporaneous papers study reconciled choices from disparate motivations. Benjamin et al. (2019) develop a survey procedure to reduce framing effects when measuring risk aversion, in order to get better risk measurements to inform retirement saving policies. They have subjects make hypothetical life-cycle savings choices which are related to one another according to various consistency principles. They allow subjects to reconsider their choices when the choices are conflicting, and ask subjects to rate whether the situations are sufficiently different to warrant choosing different options. Breig and Feldman (2019) ask subjects a number of risky budget set decisions, each repeated twice. Subjects are given the opportunity to revise
their choices, sometimes showing earlier choices on the screen. Individuals often revise, and revised choices are closer to rationality according to standard revealed preference measures. Crosetto and Gaudeul (2019) study the asymmetric dominance effect and find that the effect disappears after subjects are given time and incentives to revise their choices.

Relative to this literature, our paper allows us to show that revised choices are, indeed, more rational. In particular, revised choices correct mistakes whereby individuals violate an axiom but would have rather chosen consistently with it. Since we elicit preferences over axioms, we can show that individuals are revising their choices because they want their choices to be consistent with their preferred axioms. This avoids issues with experimenter demand effects that might be induced by showing subjects inconsistent choices without eliciting preferences over axioms. Furthermore, it allows us to more clearly interpret revised choices. In our reconciliation stage, subjects have strictly more information: in particular, they can see their axiom preferences and the implications these axioms have over their lottery choices. Revisions, then, reflect more informed preferences. Without seeing preferred axioms or having more information on preferences, it is not obvious how to interpret a subject changing their choices.

Though the environment differs, our paper is also related to the literature studying strategies in repeated games. Romero and Rosokha (2018), Cason and Mui (2019), and Dal Bó and Fréchette (2019), among others, allow subjects to design comprehensive strategies in indefinitely repeated prisoner’s dilemma games, rather than choosing actions each period. While our subjects do not design their own “axioms” to follow, our paper can be thought of as a similar procedural experiment where subjects choose rules to implement decisions for them. This is similar to the distinction between substantive and procedural rationality as outlined in Simon (1976), who calls for economists to “become interested in the procedures—the rational processes—that economic actors use to cope with uncertainty” (Simon, 1976, p.81). Halevy and Mayraz (2020) take a step in that direction, allowing subjects to design procedures to carry out their investment decisions. They find that subjects prefer using procedures to making decisions on their own, which is similar to our subjects’ preference for following the axioms.
VII. DISCUSSION

We present the first incentivized experimental evidence supporting the view that canonical choice axioms have normative content and that violations of axioms represent mistakes. In directly eliciting preferences over axioms, we find that individuals view them as rules that they want their choices to follow. When lottery choices conflict with stated axiom preferences, individuals change their choices to be consistent with the axiom, rather than inferring from their choices that the axiom is not desirable.

Our experiment takes a step towards identifying individuals’ choices as “preferences” versus “mistakes,” but also highlights the difficulties in doing so. The evidence suggests that most subjects do view these axioms as desirable and change choices accordingly, leading us to interpret their inconsistent lottery choices as mistakes. However, a substantial minority of individuals do not change their choices despite wanting to follow the axiom. In this case, it is not obvious how to declare either the axiom or lottery decisions as preferences or mistakes in these cases. However, these situations might not be surprising. For example, [Gilboa et al. (2009)](https://journals.cambridge.org/article/S0020728909000854) argue that it is natural to encounter situations where a preference for a decision rule conflicts with preferences over a single decision problem. Sometimes individuals resolve a conflict by adhering to the rule (axiom) and other times by adhering to their decisions, but neither needs to be abandoned in general. Subjects in our experiment who conflict in their rules and choices demonstrate that these cases occur in practice.

This experiment also demonstrates the difficulty of making welfare statements from choices. For example, whether or not a subject reconciles an inconsistency often depends on the question that is asked. One case of this occurs for the three questions related to independence. For question IND1, subjects are more likely to revise their choices to satisfy independence than IND2 or IND3. Thus, using independence to evaluate welfare may depend on the type of question asked. Other issues of measuring welfare are brought up in [Bernheim and Rangel (2009)](https://doi.org/10.1002/jo.20070) and focus on the difficulty of making welfare statements in the presence of framing. In contrast, all of our questions are framed the same way so we demonstrate issues of making welfare statements even holding the framing of questions fixed.

We believe our results provide implications for both theoretical models and experimental design, and we discuss these both in turn. We also see a number of open
directions for future research, and we outline these explicitly.

VII.A. Implications for Theory

Our results suggest a role for economic theory to model preferences over axioms alongside modeling the choices individuals make. Our experiment elicits two revealed preferences—one over axioms and one over lotteries. These preference relations do not always align, resulting in violations of the axioms. However, results suggest that, in these cases, the preference over axioms supersedes the lottery preference. Our results suggest that individuals do have preferences over axioms directly, so it might prove fruitful to incorporate these preferences into theoretical models. This would provide structure to exploring the interaction between axiom preferences and choices. While there is little theoretical work that explicitly models different types preferences that are related, one example is [Gilboa et al. (2010)] which models the relation between objective and subjective preferences.

Our results also contribute to an interesting discussion on the role of decision theory, outlined in [Gilboa (2010)], who writes:

“We are equipped with the phenomenally elegant classical decision theory and faced with the outpour of experimental evidence á la Kahneman and Tversky, showing that each and every axiom fails in carefully designed laboratory experiments. What should we do in face of these violations? One approach is to incorporate them into our descriptive theories, to make the latter more accurate. This is, to a large extent, the road taken by behavioral economics. Another approach is to go out and preach our classical theories, that is, to use them as normative ones... In other words, we can either bring the theory closer to reality (making the theory a better descriptive one) or bring reality closer to the theory (preaching the theory as a normative one). Which should we choose?”

Our results call for the latter, and suggest that individuals already view the classical theory as normative in many instances. For a majority of individuals, violating canonical axioms is revealed a mistake by their own choices. We help individuals make better decisions, according to their own preferences, when we assist them in satisfying these axioms. This is not to diminish the role of descriptive theories, but to draw attention to the different roles that descriptive and normative theories may play.
VII.B. Implications for Experiments

Given that we find individuals making mistakes, it’s unclear how elicited preferences would change after individuals are given the opportunity to “correct” their choices. For example, are revised choices systematically more/less risk averse than original choices? As mentioned above, we chose specific questions that would result in violations of the axioms, so our experiment was not designed to answer these questions. However, we can speculate on the implications. For example, one of our questions (TRANS3) comprised a separated price list. Price lists are common tools for experimental elicitation, and there is no general consensus as to whether experimenters should enforce a single switching point (see discussion in [Brown and Healy (2018)]). We find that over 70% of subjects who exhibit intransitivities in the form of multiple switching points change their choices to be transitive. This suggests that enforcing a single switching point might actually help subjects express their underlying desire for transitivity.

In this vein, it might be the case that there are other features of the environment or experimental design that cause axiom preferences and choice preferences to align. For example, experimental interfaces could notify subjects when they violate consistency or independence of irrelevant alternatives. If we want to elicit subjects’ “rational” preferences, then this might require more study into the structure and design of experiments.

More generally, our results suggest caution in designing and interpreting experimental tests of axioms. While it is possible to design choices to induce violations of nearly any axiom, researchers (ourselves included) should think carefully about whether these violations are revealing meaningful preferences or whether we are simply setting “traps” for subjects to fall into. If we “trick” subjects into making choices they regret, then this might not be the kind of evidence we want to use to design economic models.

VII.C. Directions for Future Research

We view our experiment as one in a new line of experiments in procedural choice. We see many interesting directions in which to take this agenda, and we outline a few below.

In our experiment, people tend to follow “rules” over following choices. It would
be interesting to understand more about when and where this is true. It is also
interesting to what aspects of the environment (e.g. framing) alters an individual’s
perceptions of decision rules. In a related study, Oprea (2019) analyzes aspects of
the decision environment that make rules more complex to implement. It would
be interesting to understand more about how these measures of complexity interact
with the questions we answer in our paper. For example, what features make axioms
more complex to understand? Are more complex axioms less appealing? Does the
complexity of the environment (here, relatively simple lottery choices) affect the rules
one wishes to implement in that environment? More generally, we believe it fruitful
to study when and why it is difficult for people to implement the principles they feel
should guide their choices.

As we referenced in the Introduction, there are many other environments in which
our methodology could prove useful. For example, researchers often have hypotheses
about competing heuristics, and could use our methodology to elicit the desirability
of these heuristics directly. In short, we could elicit what individuals “want to be” or
think they “should be” doing, in addition (or instead of) eliciting what they actually
do. Naturally these methodologies are complementary, and we believe this to be a
fruitful avenue for future research.
REFERENCES


A. Screenshots
Your choices were inconsistent with this rule. In the first decision, you chose Option A over Options B and C. In the second decision, you chose Option B over Option A. Options A and B are the same in these two decisions, so the rule would prescribe making the same choice between Options A and B in these two decisions.
Your choices were inconsistent with this rule. In the first decision, you chose Option A over Options B and C. In the second decision, you chose Option A over Option B. Options A and B are the same in these two decisions, so the rule would prescribe making different choices between Options A and B in these two decisions.
As we described in Section III, we ran an additional treatment where subjects had to pay a strictly positive cost, $1, to make decisions on their own rather than following a rule. Figure VIII shows the difference in rule selection rates across treatments. For each of the axioms, individuals are more likely to select both the axiom and the control axiom in the $1 cost treatment. The Fisher’s Exact test confirms significant differences across treatments ($p = 0.000$).

Figure IX shows the choice revisions in the $1 cost treatment. 51% of inconsistencies are revised, which is significantly lower than the 63% in our main treatment (Fisher’s Exact, $p = 0.000$). However, among reconciled choices, it is still the case that individuals reconcile in favor of the axiom. 91% of reconciliations change lottery choices, which is significantly higher than the 79% in the main treatment ($p = 0.000$)).

We find a similar story for the control axiom revisions. 47% of inconsistencies are revised, which is significantly lower than the 67% in the main treatment ($p = 0.000$). Among the revised choices, 56% are revised in favor of the axiom, which is slightly higher than the 44% in the main treatment ($p = 0.095$).

\footnote{In four instances, an individual revised lottery choices in such a way that they were still inconsistent with the axiom. We include these in the category, though results are robust to dropping these instances.}
Figure IX: Axiom Reconciliation Choices for $1 Cost Treatment

Figure X: control axiom Reconciliation Choices for $1 Cost Treatment
Overall, we find the same qualitative results in both treatments: Individuals find the axioms to be normatively appealing and revise inconsistencies in favor of the axioms. The treatment differences do give some interesting insights into how subjects perceive these axioms. Individuals are more willing to follow rules when it saves them $1, which is intuitive. They are also less likely to revise inconsistent choices. This suggests that individuals in the $0 cost treatment view selecting the axiom as indicating that it should be “always” true, while individuals in the $1 cost treatment view it as something that should be “mostly” true.

In Tables XII and XIII we report the percentage of violations and direction of reconciliation for each question.
C. Lotteries

A description of references and behavioral effects for each question can be found in Table II along with the question numbering that we use to present the results. The payment amounts span from $0 to $30, and lotteries range in expected value from $1.40 to $26. Within each question (with either two or three lotteries), the difference in expected value between the lotteries ranges from $0 to $6. On average, the expected value difference is just shy of $1.75. Furthermore, these lottery questions are not the same as the lottery questions that incentivized Block 1 rule choices. This was to ensure that subjects do not update on the type of lotteries where the rules would apply. We briefly describe the source of the questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA1–IIA3</td>
<td>Huber et al. (1982) (Attraction Affect)</td>
</tr>
<tr>
<td>IIA4</td>
<td>Huber et al. (1982) (Compromise Effect)</td>
</tr>
<tr>
<td>TRANS1–TRANS2</td>
<td>Loomes et al. (1991) (Regret Theory)</td>
</tr>
<tr>
<td>TRANS3</td>
<td>Brown and Healy (2018) (Multiple Switch Points)</td>
</tr>
<tr>
<td>IND1</td>
<td>Birnbaum and Chavez (1997)</td>
</tr>
<tr>
<td>IND2</td>
<td>Kahneman and Tversky (1979) (Certainty Effect)</td>
</tr>
<tr>
<td>IND3</td>
<td>Jain and Nielsen (2019) (Reverse Certainty Effect)</td>
</tr>
<tr>
<td>BRANCH1</td>
<td>Birnbaum and Chavez (1997)</td>
</tr>
</tbody>
</table>

**Table II: Description of Questions**

**INDEPENDENCE OF IRRELEVANT ALTERNATIVES** — We used four IIA questions similar to Huber et al. (1982). Three of the four questions targeted violations of IIA by adding a dominated lottery to a binary choice problem to “attract” the subject to the dominating alternative. We refer to these as IIA1, IIA2, and IIA3. The fourth question, IIA4, targeted a violation of IIA by adding a lottery to a binary choice problem that makes one of the initial two lotteries a “compromise” option.

**TRANSITIVITY** — We used two transitivity questions (TRANS1 and TRANS2) similar to Loomes et al. (1991) that were used to examine regret theory. In addition, we included six binary questions which together comprised a separated price list, as demonstrated in Brown and Healy (2018). These six questions involved binary comparisons between a risky lottery and a sure payment. The risky lottery was the

---

23 For example, one question was a binary choice between lottery p and $14.00. Another was the
same in all six questions while the sure payment varied. Multiple switch points on the price list constitute a violation of transitivity, which we refer to as TRANS3.

FIRST-ORDER STOCHASTIC DOMINANCE — We asked four binary questions to target violations of FOSD (FOSD1–FOSD4). All four questions followed the structure in Birnbaum and Martin (2003).

INDEPENDENCE — We included three questions that targeted violations of independence, including one question from Birnbaum and Chavez (1997) (IND1), one question from Kahneman and Tversky (1979) demonstrating the certainty effect (IND2), and one question from Jain and Nielsen (2019) demonstrating the reverse certainty effect (IND3).

BRANCH INDEPENDENCE — We included one question targeting a violation of Branch Independence from Birnbaum and Chavez (1997) (BRANCH1).

CONSISTENCY — We included two questions to target violations of consistency in which we asked two binary decision problems that were each repeated twice (CONS1 and CONS2).\footnote{We chose the binary decision problems based on the questions that were nearest to indifference in Brown and Healy (2018).}

<table>
<thead>
<tr>
<th>Question</th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Lottery C</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA1</td>
<td>60% chance of $0</td>
<td>80% chance of $0</td>
<td>80% chance of $0</td>
</tr>
<tr>
<td></td>
<td>40% chance of $6</td>
<td>20% chance of $10</td>
<td>20% chance of $7</td>
</tr>
<tr>
<td>IIA2</td>
<td>60% chance of $0</td>
<td>80% chance of $0</td>
<td>85% chance of $0</td>
</tr>
<tr>
<td></td>
<td>40% chance of $6</td>
<td>20% chance of $10</td>
<td>15% chance of $10</td>
</tr>
<tr>
<td>IIA3</td>
<td>60% chance of $0</td>
<td>80% chance of $0</td>
<td>85% chance of $0</td>
</tr>
<tr>
<td></td>
<td>40% chance of $6</td>
<td>20% chance of $10</td>
<td>15% chance of $7</td>
</tr>
<tr>
<td>IIA4</td>
<td>60% chance of $0</td>
<td>80% chance of $0</td>
<td>70% chance of $0</td>
</tr>
<tr>
<td></td>
<td>40% chance of $6</td>
<td>20% chance of $10</td>
<td>30% chance of $8</td>
</tr>
</tbody>
</table>

\[\text{choice between } p \text{ and } $14.50, \text{ another between } p \text{ and } $15.00, \text{ etc. Presenting questions of this form is a common procedure used to elicit a certainty equivalent for lottery } p.\]

\[\text{The questions are not repeated back to back.}\]
**Table IV: FOSD Questions (B FOSD A)**

<table>
<thead>
<tr>
<th>Question</th>
<th>Lottery A</th>
<th>Lottery B</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOSD1</td>
<td>10% chance of $1.25</td>
<td>5% chance of $1.25</td>
</tr>
<tr>
<td></td>
<td>5% chance of $9</td>
<td>5% chance of $1.50</td>
</tr>
<tr>
<td></td>
<td>85% chance of $9.75</td>
<td>90% chance of $9.75</td>
</tr>
<tr>
<td>FOSD2</td>
<td>10% chance of $2</td>
<td>5% chance of $2</td>
</tr>
<tr>
<td></td>
<td>5% chance of $16</td>
<td>5% chance of $3</td>
</tr>
<tr>
<td></td>
<td>85% chance of $19</td>
<td>90% chance of $19</td>
</tr>
<tr>
<td>FOSD3</td>
<td>21% chance of $1</td>
<td>1% chance of $1</td>
</tr>
<tr>
<td></td>
<td>18% chance of $10.25</td>
<td>19% chance of $2</td>
</tr>
<tr>
<td></td>
<td>61% chance of $11</td>
<td>80% chance of $11</td>
</tr>
<tr>
<td>FOSD4</td>
<td>21% chance of $0.50</td>
<td>1% chance of $0.50</td>
</tr>
<tr>
<td></td>
<td>18% chance of $13</td>
<td>19% chance of $4</td>
</tr>
<tr>
<td></td>
<td>61% chance of $16</td>
<td>80% chance of $16</td>
</tr>
</tbody>
</table>

**Table V: Transitivity Questions**

<table>
<thead>
<tr>
<th>Question</th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Lottery C</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANS1</td>
<td>30% chance of $6</td>
<td>30% chance of $0.50</td>
<td>30% chance of $8</td>
</tr>
<tr>
<td></td>
<td>30% chance of $6</td>
<td>30% chance of $11</td>
<td>30% chance of $8</td>
</tr>
<tr>
<td></td>
<td>40% chance of $20</td>
<td>40% chance of $11</td>
<td>40% chance of $8</td>
</tr>
<tr>
<td>TRANS2</td>
<td>45% chance of $7.50</td>
<td>45% chance of $1.25</td>
<td>45% chance of $9</td>
</tr>
<tr>
<td></td>
<td>25% chance of $7.50</td>
<td>25% chance of $10.50</td>
<td>25% chance of $9</td>
</tr>
<tr>
<td></td>
<td>30% chance of $19</td>
<td>30% chance of $10.50</td>
<td>30% chance of $9</td>
</tr>
</tbody>
</table>

**Table VI: Price List Transitivity Questions**

<table>
<thead>
<tr>
<th>Question</th>
<th>$A$</th>
<th>$B$</th>
<th>$C$</th>
<th>$D$</th>
<th>$E$</th>
<th>$F$</th>
<th>$G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANS3</td>
<td>25% chance of $5</td>
<td>100% chance of $14</td>
<td>100% chance of $14.50</td>
<td>100% chance of $15</td>
<td>100% chance of $15.50</td>
<td>100% chance of $16</td>
<td>100% chance of $16.50</td>
</tr>
</tbody>
</table>
### Table VII: Independence Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Lottery A</th>
<th>Lottery B</th>
<th>Lottery C</th>
<th>Lottery D</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRANCH1</td>
<td>80% chance of $0</td>
<td>80% chance of $0</td>
<td>10% chance of $2</td>
<td>10% chance of $8</td>
</tr>
<tr>
<td></td>
<td>10% chance of $2</td>
<td>10% chance of $8</td>
<td>10% chance of $12</td>
<td>10% chance of $9</td>
</tr>
<tr>
<td></td>
<td>10% chance of $12</td>
<td>10% chance of $9</td>
<td>80% chance of $15</td>
<td>80% chance of $15</td>
</tr>
<tr>
<td>IND1</td>
<td>80% chance of $0</td>
<td>80% chance of $0</td>
<td>40% chance of $0</td>
<td>40% chance of $0</td>
</tr>
<tr>
<td></td>
<td>10% chance of $6</td>
<td>10% chance of $8</td>
<td>30% chance of $6</td>
<td>30% chance of $8</td>
</tr>
<tr>
<td></td>
<td>10% chance of $11</td>
<td>10% chance of $9</td>
<td>30% chance of $11</td>
<td>30% chance of $9</td>
</tr>
<tr>
<td>IND2</td>
<td>100% chance of $10</td>
<td>20% chance of $0</td>
<td>75% chance of $0</td>
<td>80% chance of $0</td>
</tr>
<tr>
<td></td>
<td>80% chance of $13.50</td>
<td>25% chance of $10</td>
<td>20% chance of $13.50</td>
<td></td>
</tr>
<tr>
<td>IND3</td>
<td>20% chance of $10</td>
<td>100% chance of $20</td>
<td>24% chance of $10</td>
<td>8% chance of $10</td>
</tr>
<tr>
<td></td>
<td>80% chance of $30</td>
<td>8% chance of $20</td>
<td>68% chance of $30</td>
<td>8% chance of $30</td>
</tr>
</tbody>
</table>

### Table VIII: Consistency Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Lottery A</th>
<th>Lottery B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONS1</td>
<td>50% chance of $3</td>
<td>25% chance of $5</td>
</tr>
<tr>
<td></td>
<td>50% chance of $15</td>
<td>75% chance of $12</td>
</tr>
<tr>
<td>CONS2</td>
<td>50% chance of $5</td>
<td>30% chance of $0</td>
</tr>
<tr>
<td></td>
<td>50% chance of $10</td>
<td>70% chance of $15</td>
</tr>
</tbody>
</table>
D. ADDITIONAL RESULTS

<table>
<thead>
<tr>
<th>#Axioms Selected</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.7%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.9%</td>
<td>3.6%</td>
</tr>
<tr>
<td>1</td>
<td>0.9%</td>
<td>0.0%</td>
<td>0.9%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.8%</td>
</tr>
<tr>
<td>2</td>
<td>2.7%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>2.7%</td>
</tr>
<tr>
<td>3</td>
<td>3.6%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.8%</td>
<td>0.9%</td>
<td>0.0%</td>
<td>6.4%</td>
</tr>
<tr>
<td>4</td>
<td>3.6%</td>
<td>2.7%</td>
<td>1.8%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>8.2%</td>
</tr>
<tr>
<td>5</td>
<td>11.8%</td>
<td>2.7%</td>
<td>1.8%</td>
<td>0.0%</td>
<td>0.9%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>17.3%</td>
</tr>
<tr>
<td>6</td>
<td>39.1%</td>
<td>17.3%</td>
<td>0.9%</td>
<td>1.8%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.9%</td>
<td>60.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>64.5%</strong></td>
<td><strong>22.7%</strong></td>
<td><strong>5.5%</strong></td>
<td><strong>1.8%</strong></td>
<td><strong>2.7%</strong></td>
<td><strong>0.9%</strong></td>
<td><strong>1.8%</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

Table IX: Number of Axioms and control axioms Selected

Figure XI: Pairwise Correlations in Axiom Choices for Main Treatment
<table>
<thead>
<tr>
<th>Question</th>
<th># of Violators</th>
<th># Keep Inconsistent</th>
<th># Revise Axiom</th>
<th># Revise Lotteries</th>
<th># Revise Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA1</td>
<td>18</td>
<td>3</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>IIA2</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>IIA3</td>
<td>17</td>
<td>4</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>IIA4</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>FOSD1</td>
<td>59</td>
<td>27</td>
<td>11</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>FOSD2</td>
<td>55</td>
<td>25</td>
<td>12</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>FOSD3</td>
<td>48</td>
<td>26</td>
<td>11</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>FOSD4</td>
<td>32</td>
<td>17</td>
<td>7</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>TRANS1</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>TRANS2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TRANS3</td>
<td>32</td>
<td>5</td>
<td>0</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>IND1</td>
<td>22</td>
<td>7</td>
<td>4</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>IND2</td>
<td>37</td>
<td>21</td>
<td>4</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>IND3</td>
<td>37</td>
<td>17</td>
<td>7</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>BRANCH1</td>
<td>22</td>
<td>9</td>
<td>0</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>CONS1</td>
<td>19</td>
<td>7</td>
<td>4</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>CONS2</td>
<td>33</td>
<td>6</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table X:** Number of Violations and Revisions Per Question, Conditional on Selecting Axiom

<table>
<thead>
<tr>
<th>Question</th>
<th># of Violators</th>
<th># Keep Inconsistent</th>
<th># Revise Axiom</th>
<th># Revise Lotteries</th>
<th># Revise Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>c−IIA1</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>c−IIA2</td>
<td>12</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−IIA3</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c−IIA4</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD1</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD4</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>c−TRANS1</td>
<td>11</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−TRANS2</td>
<td>11</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−IND1</td>
<td>11</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>c−IND2</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>c−IND3</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−BRANCH1</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>c−CONS1</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>c−CONS2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table XI:** Number of Violations and Revisions Per Question, Conditional on Selecting control axiom
<table>
<thead>
<tr>
<th>Question</th>
<th># of Violators</th>
<th># Keep Inconsistent</th>
<th># Revise Axiom</th>
<th># Revise Lotteries</th>
<th># Revise Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA1</td>
<td>14</td>
<td>3</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>IIA2</td>
<td>21</td>
<td>7</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>IIA3</td>
<td>18</td>
<td>6</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>IIA4</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>FOSD1</td>
<td>60</td>
<td>38</td>
<td>5</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>FOSD2</td>
<td>55</td>
<td>28</td>
<td>5</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>FOSD3</td>
<td>50</td>
<td>39</td>
<td>2</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>FOSD4</td>
<td>31</td>
<td>17</td>
<td>2</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>TRANS1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TRANS2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TRANS3</td>
<td>35</td>
<td>7</td>
<td>0</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>IND1</td>
<td>21</td>
<td>7</td>
<td>1</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>IND2</td>
<td>42</td>
<td>28</td>
<td>1</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>IND3</td>
<td>37</td>
<td>24</td>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>BRANCH1</td>
<td>17</td>
<td>6</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>CONS1</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>CONS2</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table XII:** Number of Violations and Revisions Per Question, Conditional on Selecting and Violating Axiom, in $1 Cost to Decide Treatment

<table>
<thead>
<tr>
<th>Question</th>
<th># of Violators</th>
<th># Keep Inconsistent</th>
<th># Revise Axiom</th>
<th># Revise Lotteries</th>
<th># Revise Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>c−IIA1</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−IIA2</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c−IIA3</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−IIA4</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD3</td>
<td>12</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>c−FOSD4</td>
<td>15</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>c−TRANS1</td>
<td>23</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>c−TRANS2</td>
<td>24</td>
<td>14</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>c−IND1</td>
<td>29</td>
<td>21</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>c−IND2</td>
<td>17</td>
<td>11</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>c−IND3</td>
<td>17</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>c−BRANCH1</td>
<td>23</td>
<td>16</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>c−CONS1</td>
<td>14</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>c−CONS2</td>
<td>13</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table XIII:** Number of Violations and Revisions Per Question, Conditional on Selecting and Violating control axiom, in $1 Cost to Decide Treatment
To elicit willingness to reconcile inconsistent choices, we had subjects rank any of the six main axioms they selected against a $1 outside option. For example, a subject who selected all six axioms would see seven boxes on their screen—one for each of the axioms, and one with an option that says “I would rather have $1 than reconcile choices associated with any of the remaining rules.” Subjects first select the axiom they would most want to reconcile should their choices violate it, or select the outside option if they would rather have $1 than reconcile their choices. Then, subjects select the axiom they would next-most want to reconcile among the remaining axioms, and so on.

If the ranking were chosen for payment, we would randomly select two of the available axioms or the outside option and pay the subject according to the Block 4 reconciled choices from whichever axiom they ranked higher. For example, take a subject who ranks the axioms in the order FOSD > IIA > TRANS > CONS > Outside Option > BRANCH > IND. We would randomly select two options, say IIA and the Outside Option. The subject ranked IIA higher, so they would be paid based on their reconciled choices in the IIA question, as described below. Regardless of their ranking, a subject still faces all of the reconciliation decisions in Block 4; the ranking only impacts which would be paid.

Technically, the reconciled choices for the last-ranked axiom would not be incentivized in this procedure. The reconciled choices of last-ranked axiom would never be implemented since we implement the reconciled choices of whichever axiom is ranked higher. To ensure incentive compatibility, there’s an independent chance that we would randomly select the reconciled choices to pay, as described below. This means that the reconciled choices are almost twice as likely to be paid as the original choices. Therefore, if anything, subjects should be more concerned with their reconciled choices than their original choices. This also helps encourage subjects to carefully consider the reconciliation opportunity.

The main purpose of the ranking is to see whether subjects strictly prefer to re-evaluate their choices. This gives us only a coarse measure—whether individuals are

\[25\] If they had ranked the outside option higher, they would be paid for their original choices in the IIA question and would receive an extra $1 bonus. If they did not violate IIA, they would be paid for their original choices and would receive a $1 bonus, which ensures their ranking is not affected by their perceptions of which axioms they were more likely to have violated.
willing to give up at least $1. We felt this would be easier for subjects to understand than trying to elicit a finer willingness-to-pay for each axiom. We also use the reported rankings to look for any consistent patterns within rankings.\footnote{With this procedure, we cannot rule out that subjects are indifferent among the axioms or to the outside option. We look for systematic patterns in the rankings, but acknowledge this as a shortcoming of our ranking procedure. We also cannot rule out that rankings are driven by subjects’ beliefs about the expected value differences of the associated lotteries.} This gives us a finer measure of subjects’ perceptions of the rules compared to the binary ranking in Block 1.\footnote{Eliciting a ranking over axioms is also similar to the work of MacCrimmon and Larsson (1979).}

We find that, for each of the six axioms, about 60% of subjects are willing to give up at least $1 to re-evaluate their inconsistent choices. We find that 66% of subjects rank Branch Independence before the $1 outside option, and this is similar at 62% for Independence, 62% for FOSD, 62% for IIA, 60% for Transitivity, and 56% for Consistency. The pairwise comparison between Consistency and Branch Independence is significant ($p = 0.041$) but all other pairwise comparisons show insignificant differences in willingness to pay (Wilcoxon Sign Rank tests, $p > 0.134$). On an individual level, 85% of subjects are willing to give up $1 to re-evaluate choices corresponding to at least one axiom.

Looking at the average ranking of each axiom, again we find very few differences. Figure XII presents the average ranking for each of the alternatives. Transitivity and Independence are ranked lowest (most preferred), while the outside option is the least preferred, on average. There are some significant differences across axioms, but all axioms are preferred to the outside option on average. Therefore, it seems that subjects do not have any systematic preferences among the axioms they wish the follow.
**Figure XII:** Average Ranking of Each Axiom  
*Notes:* Lower ranking corresponds to the rule being more preferred.
E. SUPPLEMENTAL APPENDIX

Options: 

You Pick: 

We Pick:

\[ \text{vs. } \quad \text{vs. } \quad \text{vs. } \quad \text{vs. } \]

\[ \text{vs. } \quad \text{vs. } \]

\[ \text{Figure XIII: IIA Rule} \]

Options: 

You Pick: 

We Pick:

\[ \text{vs. } \quad \text{vs. } \]

\[ \text{vs. } \quad \text{vs. } \]

\[ \text{Figure XIV: First Order Stochastic Dominance Rule} \]

Options: 

You Pick: 

We Pick:

\[ \text{vs. } \quad \text{vs. } \quad \text{vs. } \]

\[ \text{vs. } \quad \text{vs. } \quad \text{vs. } \]

\[ \text{Figure XV: Transitivity Rule} \]
Figure XVI: Independence Rule

Figure XVII: Branch Independence Rule
Figure XVIII: Consistency Rule

Figure XIX: c-IIA Rule

Figure XX: c—First Order Stochastic Dominance Rule
Figure XXI: $c$ – Transitivity Rule

Figure XXII: $c$ – Independence Rule
Figure XXIII: $c$–Branch Independence Rule

Figure XXIV: $c$–Consistency Rule

Figure XXV: Distractor Rule #1
Figure XXVI: Distractor Rule #2

Figure XXVII: Distractor Rule #3

Figure XXVIII: Distractor Rule #4
Figure XXIX: Distractor Rule #5

Figure XXX: Distractor Rule #6
INSTRUCTIONS

This is an experiment in the economics of decision making. Stanford University and the Ohio State University have provided the funds for this research. Feel free to ask questions while we go over the instructions. Please do not speak with any other participants during the experiment and please put away your cell phones and anything that you might have brought with you.

This experiment has three different parts, and each part has many decisions. These instructions are for the first part. You will be paid based on your choice in exactly ONE randomly selected decision from the entire experiment. Each of your decisions is equally likely to be paid. In addition, you will receive $7 for completing the experiment.

For each part, we will explain how you would be paid from a decision in that part. The decision chosen to determine your payment will be shown at the end of the experiment, and we will roll dice at the front of the room to determine which decision it will be.

RULES

In this part of the experiment, you will be presented with various decision “rules.” These rules are abstract ways to represent choices, and we are interested in how people perceive them. Your task is to evaluate various rules and decide whether you would want to implement choices using the rule or make choices yourself. If you select your preferred rules carefully, will implement your preferred choices from the rules and you won’t have to make these choices on your own.

The rules will use colors to represent positive money payments. These payments could be “lotteries” or sure payments, which we’ll describe below. The possible payments range from $0 to $30. You will choose whether you want the rule to make decisions for you or whether you want to make the choices yourself. Remember, when you evaluate these rules, you should only be considering money payments, and only positive amounts (meaning you can’t lose money).

A “rule” is like an algorithm that observes your initial choices and then figures out what to pick later based on those initial choices. You can choose to use various rules to make your choices, or you can choose to make each later choice on your own.

For example, here’s one possible rule:

<table>
<thead>
<tr>
<th>Options:</th>
<th>You Choose:</th>
<th>We Choose:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="black.png" alt="Black" /> vs. <img src="white.png" alt="White" /></td>
<td><img src="black.png" alt="Black" /></td>
<td></td>
</tr>
<tr>
<td><img src="black.png" alt="Black" /> vs. <img src="gray.png" alt="Gray" /></td>
<td></td>
<td><img src="black.png" alt="Black" /></td>
</tr>
</tbody>
</table>
This rule establishes that, when given the choice between “black” and “white,” you chose “black.” Then, the rule says that if you were given the choice between “black” and “grey,” we would choose black for you.

The colors don’t have any inherent meaning (e.g. grey is not “in between” black and white). They can represent any possible money amount or lottery. As a result, you probably wouldn’t want to select this rule. If you like black over white, whatever they may be, this doesn’t necessarily mean you’d like black over grey, since grey can represent any possible payment. Grey might be $30 and black might be $5!

You want to select rules that would be **always** best for you, regardless of what the colors stand for. It’s worthwhile to make the choice on your own if the rule would not always be best for you, like in the example above. This way, if “grey” were $30, you could choose grey over black. If instead grey were $0, you’d be able to choose black.

### MIXTURES

Some of the rules will involve “mixtures.” Here’s an example.

<table>
<thead>
<tr>
<th>Options:</th>
<th>You Choose:</th>
<th>We Choose:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="black.png" alt="Black" /> vs. <img src="grey.png" alt="Grey" /></td>
<td><img src="black.png" alt="Black" /></td>
<td><img src="black.png" alt="Black" /></td>
</tr>
<tr>
<td><img src="black.png" alt="Black" /> <img src="white.png" alt="White" /> vs. <img src="grey.png" alt="Grey" /> <img src="white.png" alt="White" /></td>
<td><img src="black.png" alt="Black" /> <img src="white.png" alt="White" /></td>
<td><img src="black.png" alt="Black" /> <img src="white.png" alt="White" /></td>
</tr>
</tbody>
</table>

This rule says that if you prefer black over grey, then you would prefer a mix of black and white to the same mix of grey and white.

You can think of the mixtures like a spinner wheel. If the spinner wheel lands in the “white” region, you’d receive white. If it lands in the black (or grey) region, you’d receive black (or grey). Essentially, this rule would say that if you like black over grey, then you’d also like black-mixed-with-white over grey-mixed-with-white, regardless of what black, white, and grey are.

For example, imagine black and grey are both lotteries (remember, they can be any lottery), and imagine white is $10. If the spinner lands in the white region (which happens with probability ¼), you would receive $10. If it lands in the black or grey region (which happens with probability ¾), you would receive the original black or grey lottery.
Here’s another example:

<table>
<thead>
<tr>
<th>Options:</th>
<th>You Choose:</th>
<th>We Choose:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Black circle vs. Grey circle" /></td>
<td><img src="image2" alt="Black circle" /></td>
<td><img src="image3" alt="Grey circle with 20% chance of Black circle" /></td>
</tr>
</tbody>
</table>

Many people would find this to be a desirable rule. If you like black more than grey, you might prefer the mix of black-and-grey to grey. Because there’s a chance you can get black, which you prefer to grey.

## TYPES OF QUESTIONS

In the experiment, the colored circles will represent various lotteries. These lotteries can have any payments from $0 to $30, and can have any probabilities. For example, maybe one lottery is a 100% chance of $20. Another lottery might be a 40% chance of $5, 30% chance of $11, and 30% chance of $16.

You won’t know what the exact lotteries are while you’re making your decisions about the rules, but the lotteries are sufficiently varied. Some might be very risky, some are safe, some involve low payments, some high, etc.

Therefore, when choosing your rules, you’ll want to select only rules which are always best for you. By selecting rules which are always best for you, you’ll get the option which is best for you without having to make the decision on your own. By not selecting rules which aren’t always best for you, you’ll ensure that you have the opportunity to pick your most preferred option in the cases where the rule wouldn’t be good for you.

## PAYMENT

Along with all decisions you make in this experiment, each of your rule decisions is equally likely to be paid, and we will roll dice at the end of the experiment to determine which decision will be paid. If one of the rules is the randomly selected question to determine your payment, here’s how we would pay you.

Each rule corresponds to some specific lottery questions. If you selected to make the choice on your own, you will be shown the lottery question(s) and you will make the choice on your own.

If you selected to implement your choices using the rule, we will make the choice for you based on what the rule says.
You will be paid based on the lottery chosen, either by you or by the rule.

For example, imagine you are being paid based on the following rule:

<table>
<thead>
<tr>
<th>Options:</th>
<th>You Choose:</th>
<th>We Choose:</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Circle] vs. ![Gray Circle]</td>
<td>![Circle]</td>
<td></td>
</tr>
<tr>
<td>![Black Circle] + ![White Circle] vs. ![Gray Circle]</td>
<td>![Black Circle]</td>
<td>+ ![White Circle]</td>
</tr>
</tbody>
</table>

There would be two lotteries corresponding to “black” and “grey.” Remember, these lotteries could be anything, but let’s say the two lotteries were

1. 100% chance of $12
2. 75% chance of $20, 25% chance of $10

Let’s say you prefer option 2 (75% chance of $20, 25% chance of $10), so think of this as “black” and the other (100% chance of $12) as “grey.” In this case, you preferred black over grey, as the rule assumes.

Then, the rule says we will pay you “black & white” instead of “grey.” The white lottery also could be anything, say it’s 100% chance of $5. Then we would pay you “black & white,” or 75% chance of $20, 25% chance of $10, as well as 100% chance of $5, instead of “grey,” which would have been 100% chance of $12.

If you hadn’t selected this rule, you would make the choice of black + white vs. grey on your own.

---

**HOW TO ANSWER**

There are no “right” or “wrong” answers in these questions. We are interested in people’s preferences over these rules, so we simply want to know which rules you like and which you don’t.

Since we want to know people’s true preferences, the payment is set up such that you’ll be paid your “favorite” thing if you answer truthfully. Therefore, you have no incentive to answer differently from what you really think. If you think the rule should describe your choices, you should select it (and then you’ll save the cost of making the decision on your own). If you think there are situations where the rule would not give you your favorite option, you should not select it.
In this part of experiment, you will be making choices between “lotteries.” A lottery specifies the chance of receiving certain payoffs. In this experiment, the possible payoffs will range from $0 to $30. The chance of each payoff can be anything from 0% to 100%.

For example, one lottery could give you an 80% chance of $18, 10% chance of $7, and a 10% chance of $4. Another lottery might give a 100% chance of $13. There are many different possible lotteries.

In each decision, you will see two or three lotteries on your screen. Your screen will display the written values for the payoffs and probabilities in a table.

This is an example of what that could look like:

<table>
<thead>
<tr>
<th>Option A:</th>
<th>Option B:</th>
</tr>
</thead>
<tbody>
<tr>
<td>40% chance of $5</td>
<td>50% chance of $3</td>
</tr>
<tr>
<td>30% chance of $7</td>
<td>50% chance of $25</td>
</tr>
<tr>
<td>30% chance of $19</td>
<td></td>
</tr>
</tbody>
</table>

To visualize these lotteries, imagine rolling two 10-sided dice to generate a number from 1 to 100. In Option A, the first 40 numbers pay $5, the next 30 numbers pay $7, and the last 30 numbers pay $19. In Option B the first half of the numbers pay $3 and the other half pay $25. If this is the question to determine your payment, we would roll the dice and pay you according to which number is rolled.

Your task is simply to choose the lottery you prefer. The computer will record your choice and then will bring you to the next decision.

You will be making 33 decisions. Each decision will be presented on a different screen. Along with all decisions you make in this experiment, each of your lottery decisions is equally likely to be paid, and we will roll dice at the end of the experiment to determine which decision will be paid.

For example, in the randomly selected problem, imagine you chose the lottery which gives

| 30% chance of $0                      |
| 50% chance of $11.50                  |
| 20% chance of $17                      |

If we roll a number 1—30, you’d receive $0. There are 30 out of 100 possible numbers between 1 and 30, so this corresponds to a 30% chance of $0.

If we roll a number 31—80, you’d receive $11.50. There are 50 out of 100 possible numbers between 31 and 80, so this corresponds to a 50% chance of $11.50.
If we roll a number 81—100, you’d receive $17. There are 20 out of 100 possible numbers between 81 and 100, so this corresponds to a 20% chance of $17.

**HOW TO ANSWER**

Simply choose the lottery you prefer! In any given decision, if it’s the decision to determine your payment, we will give you the lottery you selected. So you should select the lottery that you’d rather have determine your payment.

There are no “right” or “wrong” answers in these questions. We are interested in people’s preferences over these lotteries, so we simply want to know which lotteries you prefer.

These lotteries are not the same questions as the lotteries which would determine your payment from your “rule” decisions.
RULES AND CHOICES

In the first part of the experiment, you indicated which “rules” you liked and which ones you did not like. In the second part of the experiment, you made choices over lotteries. It’s possible that you selected a rule in Part 1 which could have been applied to make one of your choices in Part 2.

In this final part of the experiment, you will have the opportunity to reevaluate your choices. We will show you if you ever selected a rule and then made lottery choices which were not what the rule would have chosen for you. This way, you can analyze the rule and the choices together. You might decide that you want to change your lottery choices to align with what the rule would have chosen for you. Or you might decide that the rule isn’t always true of your choices, so you can “un-select” it. Or, you can leave your choices as they were.

Here’s how that will work.

RANKING OF THE RULES

First, for some of the rules you selected in Part 1, we will ask you to “rank” them in order of how much you would like to revise your choices if they conflict with the rule. An example is shown on the board. Essentially, we are asking you to rank the rules according to how important you think they are. The rule you think is most important would be ranked first, the next most important would be ranked second, etc.

At the bottom of the screen, there is also an option which says “I would rather have $1 than revise my choices for any of the remaining rules.” This is an alternative which can enter into your ranking, too.

For example, let’s say you selected six rules. In order of importance, you rank them

Rule 4, Rule 1, Rule 2, Rule 3, Rule 6, Rule 5.

Now suppose that you’d be willing to pay at least $1 to revise choices that conflict with rules 4, 1, 2, and 3, but you’d rather have an extra $1 than revise choices that conflict with rules 6 and 5.

Then, you would report a ranking of

Rule 4, Rule 1, Rule 2, Rule 3, $1, Rule 6, Rule 5.

If this ranking question is randomly chosen for payment at the end of the experiment, we will randomly select two of the things you ranked. We would pay you according to whichever you ranked higher. If that is one of the rules, we would pay you for your revised choices in that rule (more on this below). If that is the $1, you would receive a $1 bonus and would be paid for your original choices.

Your decision will be implemented only IF your choices are not in agreement with the rule. If your choices are in agreement with the rule, we will just implement your original decisions and you still get a $1 bonus. This means that you don’t need to worry about whether you think your choices were in agreement with the rule or not. If they were, your decision in these questions won’t matter (since we’ll
just implement your original choices). So you should report your ranking as if your choices conflict with the rules.

---

### REVISING YOUR CHOICES

Regardless of your ranking, we will show you your rule/lottery decisions and you can choose to revise them or not. These decisions actually could be paid if you are paid for your revised choices, as described above.

To give you the opportunity to revise your choices, we will show you the rule on your screen and we’ll show you your choices which are inconsistent with that rule. An example is shown on the board.

The choices you made in Parts 1 and 2 will be shown in red. Below that, you’ll see an explanation of why the rule would have selected something different than what you chose. To change your choice of the rule, just click the box containing the rule. It will change from red to grey and vice versa. To change your lottery decisions, just click the box of the option you prefer instead. It will turn red, also. You can change any of your choices as many or as few times as you wish. Your final choices will be those in red.

You can revise your choices in any way you wish. For example, your choices being inconsistent with the rule might indicate that you don’t like the rule after all and it is not a rule which is always best for you. In that case, you could choose to “un-select” the rule. On the other hand, you might decide that the rule does give you things which are always best for you, and therefore you want to change your choices to align with the rule. Finally, you might decide that you like the rule but also like your choices, even though they’re inconsistent. In this case, you can leave all your choices as they are.

In other words, you can change any of your choices or keep them the same. You are under no obligation to make your choices agree with the rules, or vice versa.

---

### PAYMENT

If one of your revised choices is selected for payment, we would pay you in the same way that we described in Parts 1 and 2. That is, if a revised rule were selected for payment, we would pay you just like how we described for the rules in Part 1. If you selected the rule in your revised choices, we would use the rule to make decisions for you. If you did not select the rule, you would make the decision for yourself.

If a revised lottery decision were selected for payment, we would pay you the lottery you selected in your revised choices. We would roll dice to generate a number between 1—100 to determine which payoff you receive.

---

### HOW TO ANSWER

There are no “right” or “wrong” answers in these questions. We are interested in people’s preferences over these rules and choices, so we want to know which options you prefer.
Since we want to know people’s true preferences, the payment is set up such that you’ll be paid your “favorite” thing if you answer truthfully. Therefore, you have no incentive to answer different from what you really think. If you think the rule should describe your choices, you should select it. If you think there are situations where it should not describe your choices, you should not. If you like one lottery over the other, you should choose that.

Your revised rule selections or your revised lottery choices could be selected to determine your final payment. If this is the case, we will pay you based on your revised rule or lottery choice, as described in Parts 1 and 2. Therefore, since your revised choices might be selected to determine your payment, you should choose the options you most prefer.

You are also under no obligation to make your lottery choices and rules agree. And, if you do want to change your decisions, you can change either the rule or the lottery choices depending on your preferences.