

Reduction in Belief Elicitation*

Andrew Dustan¹, Kristine Koutout² and Greg Leo³

¹Department of Economics, William & Mary

²Golub Capital Social Impact Lab, Graduate School of Business,
Stanford University

³Department of Economics, Vanderbilt University

December 2023

Abstract

The state-of-the-art in eliciting probabilistic beliefs, the Binarized Quadratic Scoring Rule (BQSR), relies on an easily overlooked assumption: the reduction of compound lotteries. We find evidence that 70% of people violate the reduction assumption for at least some lotteries involved in the BQSR. Those who reveal preferences consistent with reduction are 33% more likely to report accurate beliefs. As theoretically predicted, the difference between reducers and non-reducers is attenuated in the elicitation of a 50% probability. We implement a novel Rank-Ordered Elicitation (ROE), which does not rely on reduction, to test whether eliminating the need for reduction increases the accuracy of reported beliefs. We find no evidence for this hypothesis. Taken together, the results suggest that preferences inconsistent with reduction contribute to the empirical under-performance of the BQSR, while leaving open a role for other factors in reducing behavioral incentive compatibility.

*Contact information: Leo [Corresponding], g.leo@vanderbilt.edu, Vanderbilt University Department of Economics 5201 West End Avenue Nashville, TN 37235. Dustan, addustan@wm.edu. Koutout, kkoutout@stanford.edu. The authors thank P.J. Healy, attendees of the ESA Boston 2022, and SABE Lake Tahoe 2022 for their valuable comments and feedback. Colton Cronin provided outstanding assistance in programming the experiment. This research was approved by the Institutional Review Board at Vanderbilt University March 10, 2021, with approval number 210231. This experiment was pre-registered with the AEA RCT Registry, with registry number AEARCTR-0007939.

1 Introduction

Beliefs play an important role in decision-making under uncertainty; however, beliefs are not directly observable. In economics, we generally infer beliefs from choices, but choices are determined both by beliefs and preferences. Thus, any procedure for eliciting beliefs must make important assumptions about the structure of preferences.

The current state-of-the-art in eliciting probabilistic beliefs is the binarized quadratic scoring rule (BQSR) (Hossain and Okui, 2013).¹ The common understanding of the BQSR is that it is incentive compatible for all expected utility maximizers, as well as some non-expected utility maximizers. The BQSR pays people in compound lotteries, relying on an easily overlooked assumption about preferences: the reduction of compound lotteries.

In this paper, we test empirically whether 1) this reduction assumption is violated by participants for the types of compound lotteries involved in the BQSR, 2) the failure of reduction predicts accuracy in the BQSR, and 3) a procedure that does not require reduction elicits more accurate beliefs from non-reducers. We test these pre-specified hypotheses using a simple, two-stage experiment. In the first stage, participants make choices between two compound lotteries. In the second stage, participants report their beliefs about objective probabilities. One of the compound lotteries in each of the first-stage choices corresponds to the compound lottery the participant faces *if they choose to accurately report* the objective probability elicited using the BQSR. In addition, we elicit corresponding objective probabilities using a novel procedure, the Rank-Ordered Elicitation (ROE), which does not require the reduction of compound lotteries.

We find that most participants choose compound lotteries inconsistent with reduction and that this behavior predicts inaccuracy in the BQSR. Only 30% of participants are always “reducers,” meaning that they consistently choose the compound lottery with the highest expected payoff.² In line with theory, participants whose revealed preferences are consistent with reduction in the relevant lottery are 13 percentage points more likely to report the corresponding objective probability accurately than

¹See Charness et al. (2021) for a comprehensive review of belief elicitation methods used in experimental economics.

²This does not imply that most people do not choose consistently with reduction in other settings or for other sets of compound lotteries. Our goal is to quantify the extent of reduction-consistent revealed preferences for the lotteries that participants face under the BQSR in the lab.

those who choose the compound lottery with the lower expected payoff. Unexpectedly, this difference in accuracy persists in the ROE. Indeed, the BQSR and ROE perform similarly to each other for both reducers and non-reducers. We further test for the role of reduction in the BQSR by comparing reducers to non-reducers when reporting a 50% probability. As theoretically predicted, there is no detectable difference between the performance of reducers and non-reducers in reporting the centered probability.

In summary, while most participants do not display the preferences consistent with reduction required for incentive compatibility of the BQSR, removing the need for reduction in another procedure does not improve the accuracy of reported beliefs. These results highlight the need to expand our selection criteria for belief elicitation procedures. Behavioral economics establishes the roles of attention, heuristics, and many other biases in economic decision-making. In line with Danz et al. (2022), we suggest that belief elicitation procedures should account for these behavioral factors, in addition to theoretical incentive compatibility.³

We consider elicitations of the form “what is the probability that event E occurs?” For any reported belief \tilde{p} about this probability, the BQSR results in a compound lottery. If the event occurs, the participant is paid a monetary reward with the probability $1 - (1 - \tilde{p})^2$. If the event does not occur, the participant receives the monetary reward with probability $1 - \tilde{p}^2$.

Without an assumption of reduction, there is no guarantee that the compound lottery the participant is compensated with under truth-telling is most-preferred among all other available compound lotteries. For instance, the compound lottery that results from reporting a belief of 50% pays the monetary reward with 75% probability regardless of whether the event occurs. Consider choosing between that lottery, which pays a 75% chance with a 100% chance, and the compound lottery for accurately reporting a belief of $\frac{1}{3}$: 33% chance of a 56% chance and 67% chance of an 89% chance. The latter lottery pays a higher expected simple lottery (78%), but it is reasonable to believe a participant would in fact prefer the lottery that does not correspond to accurate reporting.

To avoid the reduction assumption on compound lotteries, we develop a new procedure to elicit beliefs: the ROE. The ROE draws inspiration from multiple price lists by asking participants to rank an option in which payment is based on whether

³Danz et al. (2022) refer to this holistic criterion as “behavioral incentive compatibility.”

the event E of interest occurs relative to a list of objective lotteries that pay with known probabilities. The intuition for this procedure is that, if a participant prefers to be paid based on whether E occurs to being paid with 50% probability, then the participant believes there is a greater than 50% probability that E occurs.

We make three contributions to the literatures on reduction and belief elicitation. First, we document that a majority of people reveal preferences inconsistent with reduction of the *specific* compound lotteries used to incentivize accurate reporting in the BQSR. Concurrently, Danz et al. (2022) also find that people often do not choose the compound lottery that maximizes expected payoff in lotteries relevant to the BQSR. We make an additional contribution through our intra-participant design by providing the first evidence that people who fail to reduce are less likely to report accurate beliefs than those who reduce. This difference is mitigated in the elicitation of a centered probability (50%), affirming the importance of the reduction assumption in the BQSR. We then show that addressing the failure of the reduction assumption theoretically using the ROE does not increase accurate reporting.

A sizable empirical literature explores the plausibility of the axiom of the reduction of compound lotteries, dating at least to Bar-Hillel (1973). As Hajimoladarvish (2018) summarizes, reduction has been tested by comparing the choices between compound lotteries and their equivalent simple lotteries (Harrison et al., 2015), or by eliciting and comparing the certainty equivalents for both compound and simple lotteries (Abdellaoui et al., 2015). Support for the axiom is mixed, with findings that appear to be sensitive to framing (Bernasconi and Bernhofer, 2020), details of the experimental design (Harrison et al., 2015), and the comparison method (Hajimoladarvish, 2018). Our finding that many participants behave inconsistently with the reduction of compound lotteries is therefore not in conflict with existing literature, although it is the first to evaluate reduction as it relates to accurate reporting in the BQSR.

The use of probability currency to induce incentive compatibility in belief elicitation for people with non-risk-neutral preferences dates to the 1960s (Smith, 1961; Roth and Malouf, 1979). Hossain and Okui (2013) were the first to generalize the method of binarizing *any* proper scoring rule to achieve incentive compatibility for a wide range of preferences. Since then, the BQSR has been used in many applications (see, for example, Babcock et al., 2017 or Dianat et al., 2022). Recently, Danz et al. (2022) and Healy and Kagel (2022) identify issues with eliciting beliefs, raising questions about the empirical performance of the BQSR. We test one explanation for less

than theoretically-optimal performance of the BQSR: preferences inconsistent with the reduction of compound lotteries.

The remainder of the paper is structured as follows. Section 2 gives a theoretical treatment of the BQSR and ROE, showing that the BQSR requires reduction for incentive compatibility while the ROE does not. Section 3 presents the experimental design. Section 4 reports the main findings and the results of pre-specified hypotheses. Section 5 concludes.

2 Theory

In this section, we first provide the precise preference assumptions under which the Binarized Quadratic Scoring Rule (BQSR) is incentive compatible. Then, we present the Rank-Ordered Elicitation (ROE) for eliciting beliefs, which builds on the use of multiple price lists. Lastly, we show that the ROE requires a weaker assumption than reduction to be incentive compatible.

2.1 Binarized Quadratic Scoring Rule

The BQSR is part of a longer tradition in experimental economics of using probabilistic currency to linearize preferences. As shown in Hossain and Okui (2013), binarizing any proper scoring rule results in a procedure that yields incentive compatible elicitation of the target statistic under any expected utility preferences (regardless of risk preferences) and some non-expected utility preferences.

The stated key assumption in Hossain and Okui is monotonicity over binary objective lotteries. That is, a participant prefers a simple lottery over outcomes a and b to another simple lottery over those outcomes if and only if it yields the better outcome more often. Formally, let $L(p)$ be the binary lottery that yields outcome a with probability p and outcome b with probability $1 - p$ and assume $a \succ b$. The key assumption is that preferences over simple lotteries can be represented by a preference functional $V(\cdot)$ with the property that $V(L(p)) > V(L(p^*)) \Leftrightarrow p > p^*$.

Theorem 1 of Hossain and Okui states that under suitable requirements on the underlying scoring rule, the optimal action taken by a participant meeting this key assumption in the binarized version of the scoring rule will be the same as the action taken by a risk-neutral participant when incentivized by the standard non-binarized

version of the rule. Thus, any scoring rule that is incentive compatible for a risk-neutral participant can be extended through binarization to be incentive compatible regardless of risk preferences.

Although the monotonicity assumption on binary objective lotteries is key to the incentive compatibility of the BSR, it is not sufficient. As Hossain and Okui note, they “also assume that the agent’s preference under the BSR depends only on the probability of winning.” Implicit in this assumption is that a participant’s preferences over the *compound lotteries* that yield outcome a with *expected* probability p are the same as their preferences over *simple lotteries* that yield outcome a with probability p . To demonstrate the need for reduction in BSR more concretely, we formalize our analysis by first presenting some notation.

2.1.1 Environment

We assume that participants make choices on the basis of well-defined subjective probabilities about events, but do not necessarily have preferences consistent with expected utility. The environment follows some of the conventions in Machina (1995). Our analysis of the BQSR makes use of the following objects and types of lotteries:

The set of *outcomes* is $\mathcal{X} = \{a, b\}$.⁴ The set of *events* is $E \in 2^{\mathcal{S}}$, in which $\mathcal{S} = \{\dots, s, \dots\}$ are states of the world. *Simple Lotteries* are objective lotteries of the form $S = (a \circ p, b \circ (1 - p))$, where $p \in [0, 1]$. *Compound Lotteries* are objective lotteries of the form $C = (S_1 \circ p, S_2 \circ (1 - p))$, where $p \in [0, 1]$, and S_1, S_2 are simple lotteries. *Compound Subjective Lotteries* mix objective probabilities and subjective events and are of the form $H = (S_1 \circ E, S_2 \circ E^c)$ where S_1, S_2 are simple lotteries, E is an event and E^c is the complement of that event. Finally, the set of all *Compound Subjective Lotteries* is \mathcal{H} .

2.1.2 Incentive Compatibility and the BQSR

The BQSR asks participants to choose among a family of compound subjective lotteries indexed by probability \tilde{p} (the “reported” probability). These lotteries have the following form:

$$H(\tilde{p}) = (S(\tilde{p}) \circ E, S(1 - \tilde{p}) \circ E^c),$$

⁴Generally, the set of outcomes is assumed to be *some* finite set, but since both of the procedures discussed in this paper use only two outcomes we have included this restriction.

where the simple lottery $S(\tilde{p})$ is:

$$S(\tilde{p}) = (a \circ 1 - (1 - \tilde{p})^2, b \circ (1 - \tilde{p})^2)$$

Let p^* be the participants “true” belief about the probability of some event. Incentive compatibility requires $\forall \tilde{p} \in [0, 1]/p^* H(p^*) \succ H(\tilde{p})$. As in Hossain and Okui (2013), the only dominance assumption we will place on lottery is on simple lotteries. However, we will only be able to compare two lotteries using this dominance assumption if those lotteries can be mapped into simple lotteries.

Thus, a proof of incentive compatibility will require some way of mapping subjective compound lotteries into simple (objective) lotteries. This proof requires two assumptions: First, an assumption to eliminate the subjectivity; second, an assumption to eliminate compounding.

With this overview, we can formally state the axioms needed for the incentive compatibility of the BQSR. Let \succsim be a complete, reflexive, and transitive preference relation over \mathcal{H} . The assumptions we make of \succsim are:

Axiom 1: *Monotone on Simple Lotteries:*

$$\text{For } a \succ b, p > p' \Leftrightarrow (a \circ p, b \circ (1 - p)) \succ (a \circ p', b \circ (1 - p')).$$

Axiom 2: *Subjective/Objective Replacement on Simple Lotteries:*

$$\exists p \in [0, 1] \text{ s.t. } [S_1 \circ E, S_2 \circ E^c] \sim (S_1 \circ p, S_2 \circ (1 - p)).$$

Axiom 3: *Reduction:*

For any $p_1, p_2, \pi \in [0, 1]$

$$(a \circ p_1, b \circ 1 - p_1) \circ \pi, (a \circ p_2, b \circ 1 - p_2) \circ (1 - \pi) \sim a \circ [p_1\pi + p_2(1 - \pi)], b \circ [(1 - p_1)\pi + (1 - p_2)(1 - \pi)].$$

We can now prove incentive compatibility.

Proof: Where p^* is the participant’s true belief about the probability of event E , incentive compatibility requires:

$$\forall \tilde{p} \in [0, 1]/p^* (S(p^*) \circ E, S(1 - p^*) \circ E^c) \succ (S(\tilde{p}) \circ E, S(1 - \tilde{p}) \circ E^c)$$

By Axiom 2, we can replace the subjective lottery contingent on event E with an

objective lottery that contingent on objective probability p^* , the participants belief about the probability of E :

$$(S(\tilde{p}) \circ E, S(1 - \tilde{p}) \circ E^c) \sim (S(\tilde{p}) \circ p^*, S(1 - \tilde{p}) \circ (1 - p^*))$$

This indifference can be applied to both sides of the incentive compatibility condition. Since preferences are transitive, this results in a characterization of incentive compatibility involving only objective probabilities:

$$\forall \tilde{p} : \in [0, 1]/p^* \quad (S(p^*) \circ p^*, S(1 - p^*) \circ (1 - p^*)) \succ (S(\tilde{p}) \circ p^*, S(1 - \tilde{p}) \circ (1 - p^*))$$

Applying Axiom 3, these compound lotteries can be replaced with their induced simple lotteries. Again, by transitivity, this generates an incentive compatibility characterization involving only simple lotteries to which the dominance axiom can be applied:

$$\begin{aligned} \forall \tilde{p} : \in [0, 1]/p^* \quad & (a \circ p^* (1 - (1 - p^*)^2) + (1 - p^*) (1 - p^{*2}), b \circ p^* (1 - p^*)^2 + (1 - p^*) p^{*2}) \\ & \succ (a \circ p^* (1 - (1 - \tilde{p})^2) + (1 - p^*) (1 - \tilde{p}^2), b \circ p^* (1 - \tilde{p})^2 + (1 - p^*) \tilde{p}^2) \end{aligned}$$

Applying Axiom 1, the condition is true if and only if:

$$p^* (1 - (1 - p^*)^2) + (1 - p^*) (1 - p^{*2}) > p^* (1 - (1 - \tilde{p})^2) + (1 - p^*) (1 - \tilde{p}^2)$$

Since the right side is strictly quasi-concave and maximized at $\tilde{p} = p$, the procedure is incentive compatible.

2.1.3 Why might the incentive compatibility of the BQSR fail?

It is easy to construct preferences for which incentive compatibility fails even when the participant's preferences are monotone on binary objective lotteries (Axiom 1). These constructions involve a failure either to replace subjective events with objective probabilities (Axiom 2), or to reduce compound to simple lotteries (Axiom 3).

Every belief elicitation procedure that attempts to map subjective beliefs into an objective probability will require a comparison of subjective and objective lotteries. Thus, a failure of Axiom 2 (or some related assumption) would cause problems for any probabilistic belief elicitation. Thus, we focus on the types of preference that might violate the reduction axiom (Axiom 3).

Example: A participant believes the probability of event E is $\frac{1}{3}$. The participant is indifferent between \$10 if event E occurs and \$10 with a $\frac{1}{3}$ chance. However, the participant strictly prefers:

$$(\$10 \circ 0.75, \$0 \circ 0.25) \circ \frac{2}{3}, (\$10 \circ 0.75, \$0 \circ 0.25) \circ \frac{1}{3}$$

to

$$(\$10 \circ 0.89, \$0 \circ 0.11) \circ \frac{2}{3}, (\$10 \circ 0.56, \$0 \circ 0.44) \circ \frac{1}{3}$$

These are precisely the compound lotteries that result from reporting respectively 50% and 33.3% in the BQSR for this participant once the subjective events are replaced with the objective probabilities associated with the participant’s beliefs about those events. These preferences are inconsistent with reduction and would lead the participant to prefer reporting a 50% belief in the BQSR to reporting their true belief of (approximately) 33.3%.⁵

An extreme form of this failure in reduction would occur if a participant’s preferences over compound lotteries are such that they maximize the minimum probability of the better outcome over all contingencies in a compound lottery. In fact, this would lead a participant to *always* report 50% in the BQSR.

The preferences in this example are consistent with a form of aversion to variance *within* compound lotteries.⁶ On the other hand, the ROE does not rely on the kind of compound lotteries which are problematic in these examples. Specifically, reduction is not required.

2.2 Rank-Ordered Elicitation

The Rank-Ordered Elicitation (ROE) builds on a multiple price list methodology. Informally, it draws on the following observation: if a participant prefers to be paid, for instance, a with a 50% chance to being paid a if the event E occurs, but prefers to be paid if E happens to being paid with a probability 40%, then their belief about the probability that E occurs must be within the range 40% to 50%.

⁵We note that the reduced forms of these lotteries are not that different. They reduce to 0.78 and 0.75 probabilities of winning \$10, respectively. These weak incentives are a general problem for eliciting probabilistic beliefs with a scoring rule.

⁶This failure could be responsible for the pull-to-center effect found in Danz et al. (2022).

2.2.1 Description of ROE

While the BQSR “pays” in *compound subjective lotteries*, the ROE pays in what we call *simple mixtures*, which are probability distributions over *simple (objective) lotteries* and *simple subjective lotteries* (a type of lottery that does not appear in the proof of incentive compatibility for the BQSR).

We define each of these as follows: *Simple Subjective Lotteries* are lotteries of the form $S_E = (a \circ E, b \circ E^c)$ where E is an event and E^c is the complement of that event; *Simple Mixtures* of the form $M = (L_1 \circ p_1, L_2 \circ p_2, \dots, L_n \circ p_n)$ where $p_i \in [0, 1]$ with $\sum_{i=1}^n p_i = 1$, and each L_i is either a simple lottery or a simple subjective lottery; *The set of all Simple Mixtures* is \mathcal{M} .

Formally, to elicit the probability of event E , construct a set of m simple objective lotteries $\{S_{p_i}\}_{i=1}^m$ of the form $S_{p_i} = p_i \circ a, (1 - p_i) \circ b$ with $p_i > p_{i+1}$. In addition, construct a single simple subjective lottery S_E for the event E in question.

Eliciting a participant’s belief about the event E amounts to determining the position of the lottery S_E in their linear preference ordering among the objective lotteries $\{S_{p_i}\}_{i=1}^m$.

To elicit the position that this lottery falls in the linear order, ask the participant to rank S_E in their preferred position among the ordered simple objective lotteries $\{S_{p_i}\}_{i=1}^m$. To incentivize truthful ranking, pay the participant by randomly choosing one objective lottery S_{p_i} and paying according to whichever lottery the participant ranks higher: the objective lottery S_{p_i} or the subjective lottery S_E . In other words, if the participant ranks S_{p_i} above S_E , pay according to S_{p_i} . Otherwise, pay according to S_E . An example is given in Table 1.

This procedure produces simple mixtures of the form: $(L_1 \circ \frac{1}{m}, \dots, L_m \circ \frac{1}{m})$ where L_i is whichever lottery from each pair S_{p_i}, S_E that the participant ranks higher.

Subjective Lottery (H)	Objective Lotteries (R_{p_i})
\$10 if Event is True	0% chance of \$10
	20% chance of \$10
	40% chance of \$10
	60% chance of \$10
	80% chance of \$10
	100% chance of \$10

Table 1: ROE Example

The participant chooses where to rank the subjective lottery on the left among the objective lotteries on the right.

2.2.2 Incentive Compatibility and the ROE

First, we state the axioms needed for the incentive compatibility of our procedure. Let \succsim be a complete, reflexive, and transitive preference relation over \mathcal{M} . The assumptions we make of \succsim are:

Axiom 1: *Monotone on Binary Objective Lotteries:*

For $a \succ b$, $p > p' \Leftrightarrow (a \circ p, b \circ (1 - p)) \succ (a \circ p', b \circ (1 - p'))$.

Axiom 2': *Subjective/Objective Replacement on Degenerate Simple Lotteries:*⁷

$\exists p \in [0, 1]$ s.t. $[a \circ E, b \circ E^c] \sim (a \circ p, b \circ (1 - p))$

Axiom 3': *Statewise Monotone on Simple Mixtures:*

$L_i^* \succ L_i \Leftrightarrow (L_1 \circ p_1, \dots, L_i^* \circ p_i, \dots, L_n \circ p_n) \succ (L_1 \circ p_1, \dots, L_i \circ p_i, \dots, L_n \circ p_n)$

We can now prove incentive compatibility of the ROE under these axioms.

Proof: By Axiom 1, $\{S_{p_i}\}_{i=1}^m = S_{p_1} \succ S_{p_2} \succ \dots \succ S_{p_m}$ is a linear order. Let p^*

⁷Axiom 2' is a weakening of 2.

be the participant's belief about event E . That is, p^* solves $S_E \sim (a \circ p, b \circ (1 - p))$ according to Axiom 2.

Denote this objective lottery by S_{p^*} . By Axiom 1, S_{p^*} is somewhere in the linear order of $\{S_{p_i}\}_{i=1}^m$. By transitivity, S_E takes the same place in this linear order. Thus, the participant's preference relation on $\{\{S_{p_i}\}_{i=1}^m, S_E\}$ is a linear order. Let i^* be the smallest p_i such that $S_{p_i} \succ S_E$, and let $i = 0$ if $S_E \succ S_{p_1}$. We call i^* the participant's belief type, in which a participant of belief type i^* has a belief $p^* \in [p_{i^*}, p_{i^*+1}]$. There are $n + 1$ possible types $i^* \in (0, \dots, n)$.

Let L_i be the simple mixture that results from ranking S_E between S_{p_i} and $S_{p_{i+1}}$. This is the mixture:

$$L_i = \left(S_{p_1} \circ \frac{1}{m}, \dots, S_{p_i} \circ \frac{1}{m}, S_E \circ \frac{1}{m}, \dots, S_E \circ \frac{1}{m} \right)$$

Incentive compatibility of the ROE requires that for a player of belief type i^* , $\forall i \in \{0, \dots, n\}$: $L_{i^*} \succeq L_i$. Consider L_{i^*} and L_{i^*+1} . These lotteries are:

$$L_{i^*} = \left(S_{p_1} \circ \frac{1}{m}, \dots, S_{p_{i^*}} \circ \frac{1}{m}, S_E \circ \frac{1}{m}, \dots, S_E \circ \frac{1}{m} \right)$$

$$L_{i^*+1} = \left(S_{p_1} \circ \frac{1}{m}, \dots, S_{p_{i^*}} \circ \frac{1}{m}, S_{p_{i^*+1}} \circ \frac{1}{m}, S_E \circ \frac{1}{m}, \dots, S_E \circ \frac{1}{m} \right)$$

These lotteries differ by a single replacement of S_E with $S_{p_{i^*+1}}$. Since $S_E \succeq S_{p_{i^*+1}}$, $L_{i^*} \succeq L_{i^*+1}$ by Axiom 3'. Additional applications of Axiom 3' reveal that $L_i \succ L_{i+1}$ for $i > i^*$ and $L_i \succ L_{i-1}$ for $i < i^*$. Thus, $L_{i^*} \succeq L_i \forall i \in \{0, \dots, n\}$. The procedure is incentive compatible.

2.2.3 Comparing Statewise Monotonicity and Reduction

Statewise Monotone on Simple Mixtures (Axiom 3') and reduction (Axiom 3) are two fundamentally different preference assumptions. Axiom 3' is a dominance relation (the statements involve \succ) and Axiom 3 is a substitution axiom (the statements involve \sim). What further complicates the matter is that the types of lotteries involved in the BQSR (compound subjective lotteries) are different from the types involved in the ROE (simple mixtures). In fact, it is precisely this dichotomy that eliminates the need for reduction in the ROE.

If we were to extend the reduction axiom to assume that participants also reduce

simple mixtures, statewise monotonicity would indeed be a weaker assumption than reduction. Since reporting truthfully maximizes the probability of the good outcome in each sub-lottery in the simple mixture paid by the ROE, reporting truthfully would also maximize the expected probability. Thus, the ROE is also incentive compatible for participants who reduce simple mixtures.

On the other hand, *the reverse is not true*. That is, the BQSR may not be incentive compatible for a participant who does not reduce the compound lotteries involved in the BQSR, but who does have statewise monotonic preferences.

In addition, statewise monotonicity is required for participants to treat the incentives presented in multiple tasks independently within an experiment with random task payment (Azrieli et al., 2018). Therefore, any experiment with random task payment structure that has the BQSR as one of the tasks requires both Axioms 3 and 3'. Thus, within these experiments, the additional assumptions required for ROE to be incentive compatible are strictly weaker than those required for incentive compatibility of the BQSR.

3 Experimental Design

The goal of this experiment is to test whether reducers of compound lotteries report more accurate beliefs than non-reducers in a belief elicitation procedure that requires reduction for incentive compatibility; then, test whether this gap is reduced in a belief elicitation procedure that does not require reduction. To this end, the experiment is divided into two parts. Part one measures whether a participant chooses the compound lottery with the highest expected probability of winning. Part two elicits their beliefs using the two belief elicitation procedures. In part one of the experiment, participants complete two tasks. Each task requires a choice between two compound lotteries. One task, which we call the “1/3 lottery task,” is the choice between (1) and (2).

$$\frac{1}{3}(75\%) + \frac{2}{3}(75\%) \tag{1}$$

$$\frac{1}{3}(56\%) + \frac{2}{3}(89\%) \tag{2}$$

The other task, which we call the “5/6 lottery task,” is the choice between (3) and (4).

$$\frac{1}{6}(75\%) + \frac{5}{6}(75\%) \tag{3}$$

$$\frac{1}{6}(31\%) + \frac{5}{6}(98\%) \tag{4}$$

The order of the tasks is randomized, as well as the order (from top to bottom) of the two compound lotteries within each task.

The names of the compound lottery tasks reflect their correspondence to the lotteries the participant faces under the Binarized Quadratic Scoring Rule (BQSR) when choosing between reporting the objective probabilities being elicited in part two of the experiment or reporting 1/2. That is, a participant who reports 1/2 in the BQSR has a 75% probability of winning the prize regardless of the outcome of the event. On the other hand, if the participant reports 1/3 in the BQSR (one of the objective probabilities), they have a 1/3 chance of a 56% probability of winning and a 2/3 chance of a 89% probability of winning, corresponding to the two compound lotteries in the 1/3 lottery task. By eliciting these preferences over compound lotteries, we can say whether a participant faced with the choice of whether to accurately report the objective probability under the BQSR might not do so because there exists at least one compound lottery that they prefer over the compound lottery that offers the highest expected probability of winning.

We call the participants who choose the 75% lottery “non-reducers” in that task. We refer to the complement of this subgroup as “reducers,” but it is important to note that a participant in this subgroup could be a non-reducer when the alternative is some other compound lottery. So, reducer and non-reducer attributes are task-specific (i.e., based on whether the participant chose the lottery with the highest expected probability of winning in that task).

In part two of the lab experiment, participants complete four belief elicitation tasks. Each task is the elicitation of the participant’s belief about the probability of a set of numbers being rolled by a fair, six-sided die. Two of the belief elicitations use the BQSR and two use the Rank-Ordered Elicitation (ROE). For each procedure, the two belief elicitations are for (1) a set of numbers that correspond to a 1/2 probability (either {1,3,5} or {2,4,6}), and (2) a set of numbers that correspond to either a 1/3 ({5,6}) or a 5/6 ({1,2,3,4,5}) probability. For ease of exposition, we refer to (1) as the centered probability and (2) as the non-centered probability. The instructions for each belief elicitation procedure immediately precede the two belief elicitations for

that procedure. The order of the belief elicitation procedures, the set of numbers for the centered probability, the set of numbers for the non-centered probability, and the order in which beliefs are elicited within each procedure are block randomized.

We chose to elicit a centered and a non-centered probability because the predicted behavior of non-reducers differs between these elicitations. As we demonstrate in Appendix A, accurately reporting 50% in the centered elicitation is consistent with a wide variety of preferences that are also inconsistent with reduction.⁸ So, if reduction affects the empirical performance of the BQSR in the elicitation of the centered probability, we would expect for the difference between reducers and non-reducers to decrease in the centered elicitation.

3.1 Implementation

Participants in the experiment were recruited through Prolific.co. They were paid \$3 for completing the experiment and could earn another \$5 based on their choice in the randomly drawn task at the end of the experiment, for total potential earnings of \$8. After clicking on the link to start the experiment, participants received instructions on the general form of the experiment, their anonymity, and their payment for completing the experiment. Then the participants completed the two tasks in part one of the experiment and the four tasks in part two of the experiment for a total of six tasks. The experiment was concluded by informing participants which of the six tasks was randomly selected for payment and completing the randomization required to determine payment for that task. Screenshots of the full experiment are included in the Online Appendix. We recruited a total of 603 participants in July 2021. They took an average of 18.5 minutes to complete the survey and 73% earned the full \$8. Table 2 shows the basic demographics of the sample provided by Prolific. We did not collect demographic data as part of the experiment.

⁸Under reduction, indifference curves are linear over the pairs of sublotteries resulting from the BQSR. For the elicitation of a 50% belief, incentive compatibility extends to preferences that generate convex indifference curves over these sublotteries.

	N	Mean
Age	581	26.0 (6.03)
Man	577	0.55
Ethnicity	574	
<i>White</i>		0.59
<i>Black</i>		0.19
<i>Mixed</i>		0.13
<i>Asian</i>		0.04
<i>Other</i>		0.06

Table 2: Demographic Characteristics from Prolific

Note: Man and Ethnicity are sample proportions. Standard deviation is in parentheses.

4 Results

In this section, we first present the descriptive results of the experiment for the two types of tasks. Then, we test our three pre-specified hypotheses. Finally, we discuss and present further results.

4.1 Descriptive Statistics

The results of the compound lottery tasks in Table 3 show that participants are close to evenly split between “always reducers” (30%), “never reducers” (30%), and “sometimes reducers” (40%). Sometimes reducers predominantly choose to reduce when faced with the 5/6 lottery task and do not reduce when faced with the 1/3 lottery task.

The results of the belief elicitation tasks in Figure 1 show that both procedures elicited beliefs that are broadly consistent with the objective probabilities. The modal reported belief in each panel is the objective probability. For the non-centered probabilities, the mean of each distribution is between the objective probability and the midpoint of the support (50%).

		1/3 Lottery Task	
		No	Yes
5/6 Lottery Task	No	29.85	9.62
	Yes	30.85	29.68

Table 3: Consistency of compound lottery preferences with reduction

Note: Each cell shows the percentage of participants whose choices were (“yes”) or were not (“no”) consistent with reduction in each compound lottery task. For example, the “No/No” cell gives the percentage of participants who chose the constant 75% chance of winning the \$5 in both the 1/3 and 5/6 lottery tasks. All participants fall into one of the four cells, so that the percentages in the cells sum to 100%.

4.2 Pre-registered Hypotheses

We test three pre-specified hypotheses. The motivation for these hypotheses is that we want to learn whether the BQSR’s suboptimal empirical performance is due to a failure of the assumption that people’s preferences are consistent with reduction. We first want to know whether a procedure that does not require reduction yields more accurate reporting than the BQSR. Then, whether reducers are more likely to report accurately than non-reducers in the BQSR. Lastly, we want to learn if any difference in performance between the two belief elicitation procedures is correlated with preferences inconsistent with reduction. We use the elicitation of non-centered probabilities exclusively to test these hypotheses.

Hypothesis 1 *The proportion of participants who accurately report the objective non-centered probability is higher using the ROE compared to the BQSR.*

Hypothesis 2 *The proportion of reducers is higher than the proportion of non-reducers who accurately report the objective non-centered probability using the BQSR.*

Hypothesis 3 *The difference between the ROE and BQSR in the proportion of participants who accurately report the objective probability is larger for non-reducers than for reducers.*

Accuracy is assessed based on the relevant interval. So, when reporting the belief that a 5 or 6 is rolled by a fair, six-sided die, a participant would be counted as

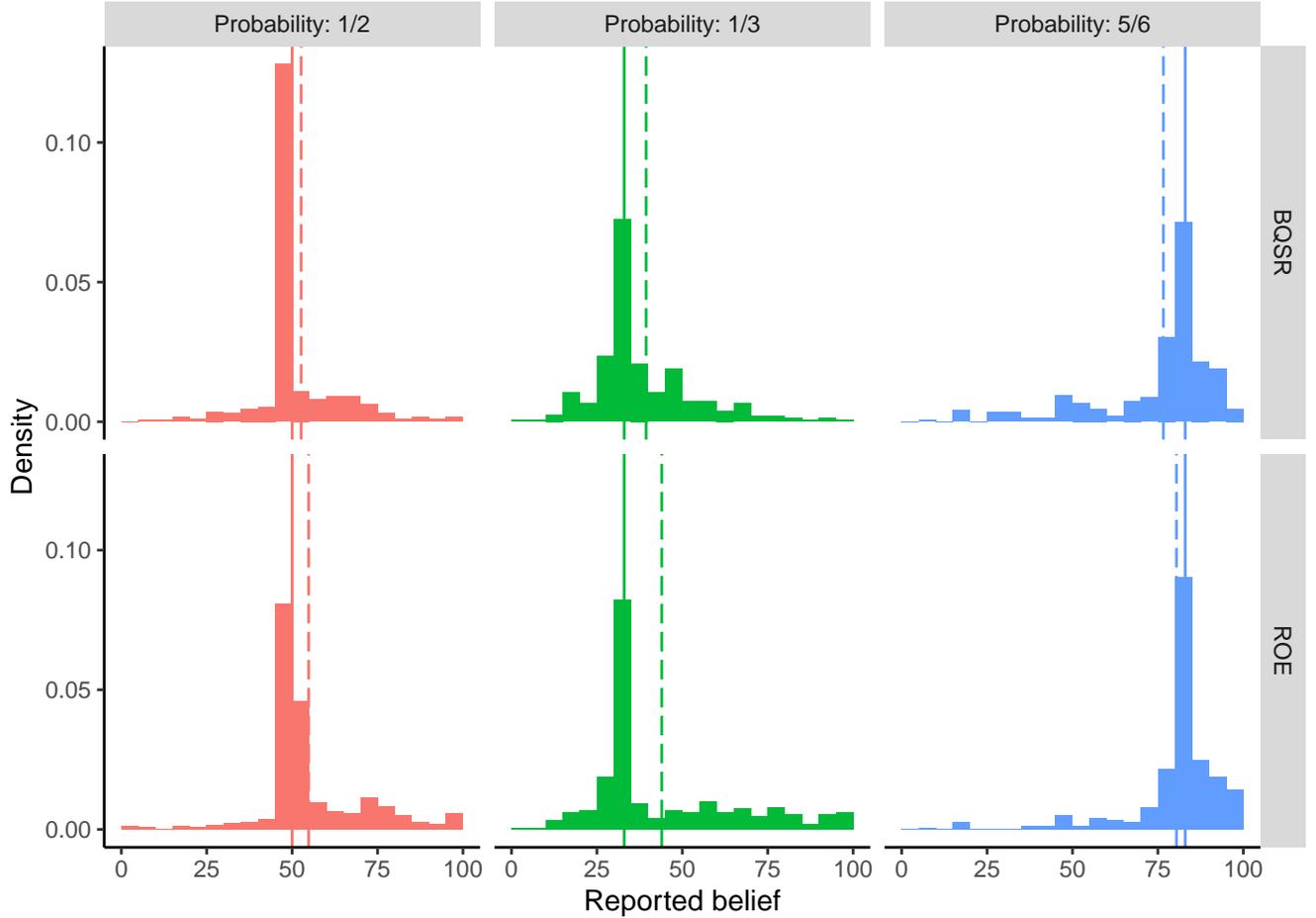


Figure 1: Distribution of reported beliefs, by elicitation procedure and objective probability

Note: Margin labels indicate the elicitation procedure and objective probability. Solid lines represent the objective probability. Dashed lines represent the mean of the reported beliefs, where the ROE beliefs are assigned to the midpoint of the reported bin.

reporting 33% accurately if they report the interval 30-35% in the ROE and report any percentage from 30-35 in the BQSR. These intervals are dictated by the 5 percentage point bins used in the ROE.

We test Hypothesis 1 by estimating:

$$accurate_{it} = \alpha + \beta_1[ROE]_{it} + \varepsilon_{it} \tag{5}$$

using data from the non-centered probabilities only, where $accurate_{it}$ is a binary

variable equal to one if participant i accurately reports in task t and $[ROE]_{it}$ is a binary variable equal to one if the belief is elicited using the ROE. Standard errors are clustered at the participant level.

For the next two hypotheses, we define participants as either reducers or non-reducers.⁹ Since the reduction of compound lotteries is only theoretically relevant to the BQSR, participants are defined based on their choice in the compound lottery task in part one of the experiment that corresponds to the non-centered belief elicited using the BQSR. That is, participants are classified as non-reducers if they were randomly assigned to report the objective probability 1/3 using the BQSR and chose the constant 75% probability in the 1/3 lottery task. If they were instead randomly assigned to report the objective probability 5/6 using the BQSR, participants are classified as non-reducers if they chose the constant 75% probability in the 5/6 lottery task.

We test Hypothesis 2 that reducers report more accurate beliefs using the BQSR than non-reducers by estimating:

$$accurate_i = \alpha + \beta_1[non-reducer]_i + \varepsilon_i \quad (6)$$

using data from the non-centered probability elicited using the BQSR only.

We test Hypothesis 3 that the difference in accuracy between the ROE and BQSR is due to non-reducers reporting less accurately in the BQSR compared to the ROE by estimating:

$$accurate_{it} = \alpha + \beta_1 1[non-reducer]_i + \beta_2 1[ROE]_{it} + \beta_3 1[non-reducer]_i * 1[ROE]_{it} + \varepsilon_{it} \quad (7)$$

using data from the non-centered probabilities only.

We find no evidence in favor of Hypothesis 1 that the ROE elicits more accurate reports than the BQSR. Regression results in column (1) of Table 4 show that the coefficient (-3.5 p.p.) is the opposite of the anticipated positive difference between the ROE and the BQSR. The 95% confidence interval bounds the positive difference at 1.1 percentage points.¹⁰ This and all other results are robust to inclusion of block-

⁹This definition is distinct from the always/sometimes/never reducers described at the beginning of this section.

¹⁰Ex ante power calculations for this sample size gave us an 80% probability of detecting an effect

randomization stratum fixed effects, as shown in Appendix Table B.1.

On the other hand, we find strong evidence in favor of Hypothesis 2 that reducers report more accurately than non-reducers in the BQSR. Regression results in column (2) of Table 4 show that the proportion of reducers who accurately report the objective probability is approximately 13 p.p. ($SE = 4.0$ p.p.) larger than the proportion of non-reducers. This difference is a 32.5% increase over the baseline 40% of non-reducers who accurately report the objective probability.

	(1)	(2)	(3)
Constant	0.466*** (0.020)	0.532*** (0.029)	0.532*** (0.029)
ROE	-0.035 (0.024)		-0.047 (0.033)
Non-reducer		-0.130*** (0.040)	-0.130*** (0.040)
Non-reducer \times ROE			0.024 (0.047)
Observations	1,206	603	1,206
Adjusted R ²	0.000	0.015	0.013

Table 4: Accuracy for reducers and non-reducers in the BQSR and ROE tasks

Note: Estimated coefficients are from linear regressions of the accuracy dummy variable on the indicated regressors. Unit of observation is the participation-elicitation pair. Sample in columns 1 and 3 is all non-centered elicitations. Sample in column 2 is non-centered BQSR elicitations only. Standard errors clustered at the participant level in columns 1 and 3 in parentheses. Heteroskedasticity-robust standard errors in column 2 in parentheses. ***: 0.01, **: 0.05, *: 0.1.

Our last pre-specified hypothesis was predicated on the positive difference hypothesis of 6.6 percentage points.

esized in Hypothesis 1; however, we found no such difference. Testing this hypothesis still allows us to learn whether there is a consistent difference between reducers and non-reducers in a procedure that does not require reduction. Regression results in column (3) of Table 4 show no difference in the differences between reducers and non-reducers when comparing the accuracy of these two subgroups in the ROE versus the BQSR. Recall that the reducer attribute is determined based on the compound lottery that corresponds to the objective non-centered probability for the BQSR. So, the difference in differences is comparing the relative performance of two fixed groups across the belief elicitation procedures.

4.3 Discussion & Further Results

From our pre-specified hypotheses, we learn that non-reducers report less accurate beliefs; however, using a procedure that theoretically eliminates the need for reduction has no effect on this difference in accuracy between reducers and non-reducers. The natural subsequent question is: Does non-reduction directly affect the performance of the BQSR, or is non-reduction proxying for another characteristic like mental acumen or inattention?

To answer this question, we compare the accuracy of reducers and non-reducers in the 50% elicitation. Recall that we choose to use the elicitation of non-centered probabilities exclusively for our pre-specified hypotheses because accurately reporting 50% in the centered elicitation is consistent with a wide variety of preferences that are inconsistent with reduction. If non-reduction predicts performance in the BQSR, but the relationship is due to a non-preference characteristic correlated with preferences for reduction, then we would expect that characteristic would similarly affect people in the elicitation of the centered probability as the non-centered probability. To test this hypothesis, we estimate the following equation using the BQSR elicitations:

$$\begin{aligned}
 \text{accurate}_{it} = \alpha + \beta_1 1[\text{non-reducer}]_i + \beta_2 1[\text{non-centered}]_{it} + \\
 \beta_3 1[\text{non-reducer}]_i * 1[\text{non-centered}]_{it} + \varepsilon_{it}
 \end{aligned}
 \tag{8}$$

If non-reducers report similarly for centered and non-centered probabilities, we would expect $\beta_3 = 0$. Instead, column 3 of Table 5 shows that we can reject that hypothesis. The gap in accuracy between reducers and non-reducers is wider in elicitations of non-

centered probabilities by 8.5 p.p. ($SE = 4.3$ p.p.). This result provides evidence that non-reduction plays a direct role in the performance of the BQSR.

	(1)	(2)	(3)
Constant	0.634*** (0.023)	0.713*** (0.018)	0.736*** (0.026)
Non-reducer	-0.088*** (0.032)		-0.045 (0.037)
Non-centered elicitation		-0.247*** (0.022)	-0.204*** (0.030)
Non-reducer \times Non-centered elicitation			-0.085** (0.043)
Observations	1,206	1,206	1,206
Adjusted R ²	0.007	0.062	0.071

Table 5: Accuracy for reducers and non-reducers in centered versus non-centered probabilities in the BQSR tasks

Note: Estimated coefficients are from linear regressions of the accuracy dummy variable on the indicated regressors. Unit of observation is the participation-elicitation pair. Sample is all BQSR elicitations, both centered (50%) and non-centered (not 50%). Standard errors clustered at the participant level in parentheses. ***: 0.01, **: 0.05, *: 0.1.

5 Conclusion

In this paper, we conduct three pre-specified tests to study how participants' revealed preferences with respect to the reduction of compound lotteries impact state-of-the-art

belief elicitation procedures. First, we show that most participants choose compound lotteries that yield a lower expected payoff compared to their alternatives, providing a preference-based explanation for why participants often report inaccurate beliefs. Second, we show that participants who choose the compound lotteries inconsistent with reduction are also less likely to report accurate beliefs compared to those who choose the lotteries with the highest expected payoffs. Our third test finds, perhaps surprisingly, that a procedure that eliminates the need for the reduction assumption does not increase accuracy. Lastly, our exploratory analysis confirms the role of reduction in the accuracy gap between reducers and non-reducers in non-centered probabilities, suggesting the importance of behavioral incentive compatibility of belief elicitation procedures.

One alternative explanation is that choices inconsistent with reduction are a mistake (Nielsen and Rehbeck, 2022) rather than truly representative of underlying non-standard preferences. Danz et al. (2022) provide some evidence that inaccurate reporting is not due to the inability to calculate the reduced lottery. They also find no correlation between accurate reporting and a cognitive reflection test. While we cannot entirely rule out that non-preference factors contribute to the inaccurate reporting we observe, the implications of preference-based and mistake-based explanations are the same: Current state-of-the-art technology in belief elicitation performs suboptimally for people who reveal preferences inconsistent with the reduction of compound lotteries.

We propose that future research focus on behavioral factors that may play a role in how participants engage with belief elicitation procedures. One potential approach is to elicit coarser beliefs, reducing the complexity of the elicitation task (Healy and Leo, 2022). Another approach would be to improve the user experience of existing procedures by simplifying the user interface and shortening the instructions. Lastly, it is an open question whether these issues with inaccurate reporting are specific to the elicitation of probabilities. The BQSR procedure could be more effective at eliciting quantiles of a distribution compared to a probability (or a mean more generally) due to the corresponding linear incentive structure.

References

- Abdellaoui, M., Klibanoff, P., and Placido, L. (2015). Experiments on compound risk in relation to simple risk and to ambiguity. *Management Science*, 61(6):1306–1322.
- Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.
- Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, 107(3):714–47.
- Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational behavior and human performance*, 9(3):396–406.
- Bernasconi, M. and Bernhofer, J. (2020). Catch me if you can: Testing the reduction of compound lotteries axiom in a tax compliance experiment. *Journal of Behavioral and Experimental Economics*, 84:101479.
- Charness, G., Gneezy, U., and Rasocho, V. (2021). Experimental methods: Eliciting beliefs. *Journal of Economic Behavior & Organization*, 189:234–256.
- Danz, D., Vesterlund, L., and Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review*.
- Dianat, A., Echenique, F., and Yariv, L. (2022). Statistical discrimination and affirmative action in the lab. *Games and Economic Behavior*, 132:41–58.
- Hajimoladarvish, N. (2018). How do people reduce compound lotteries? *Journal of Behavioral and Experimental Economics*, 75:126–133.
- Harrison, G. W., Martínez-Correa, J., and Swarthout, J. T. (2015). Reduction of compound lotteries with objective probabilities: Theory and evidence. *Journal of Economic Behavior & Organization*, 119:32–55.
- Healy, P. J. and Kagel, J. (2022). Testing elicitation mechanisms via team chat. *Working paper*.
- Healy, P. J. and Leo, G. (2022). Coarse belief elicitation. *Working paper*.

- Hossain, T. and Okui, R. (2013). The binarized scoring rule. *Review of Economic Studies*, 80(3):984–1001.
- Machina, M. J. (1995). Non-expected utility and the robustness of the classical insurance paradigm. *The Geneva Papers on Risk and Insurance Theory*, 20(1):9–50.
- Nielsen, K. and Rehbeck, J. (2022). When choices are mistakes. *American Economic Review*, 112(7):2237–68.
- Roth, A. E. and Malouf, M. W. (1979). Game-theoretic models and the role of information in bargaining. *Psychological review*, 86(6):574.
- Smith, C. A. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society: Series B (Methodological)*, 23(1):1–25.

Appendix A: Additional Theory– 50% Belief

In this section, we briefly outline the theory behind our claim that many reasonable preference models are consistent with truth-telling in the $p = 0.5$ lottery elicitation under the BQSR.

In the BQSR, regardless of belief, the participant receives a compound lottery over outcomes a and b that has two simple sublotteries. These are the lotteries that occur respectively if the event is true and if it is false. Here, as before, \tilde{p} is the stated belief.

$$L_T = (a \circ 1 - (1 - \tilde{p})^2, b \circ (1 - \tilde{p})^2)$$

$$L_F = (a \circ 1 - \tilde{p}^2, b \circ 1 - (1 - \tilde{p})^2)$$

Since these lotteries are binary over the same outcomes, they can be characterized by the probability of event a . Thus, for any \tilde{p} , L_T and L_F and thus the compound lottery above can be written as the ordered pair $(1 - (1 - \tilde{p})^2, 1 - \tilde{p}^2)$. By mapping the possible compound lotteries in the BQSR to these ordered pairs, we can apply an elementary two-good choice analogy to the participant's problem of choosing a probability in the BQSR.

Letting p_t be the probability of outcome a if the event is true and p_f be the probability of outcome a if the event is false, the budget is given explicitly by:

$$p_f = p_t + 2\sqrt{1 - p_t} - 1$$

This is a strictly concave budget symmetric about the 45-degree line and including endpoints $(1, 0)$, $(0, 1)$ as well as the point $(0.75, 0.75)$. This is the budget faced by participants in the BQSR regardless of belief. Preferences over these bundles (p_t, p_f) depend on the participants actual belief as well as the details of their preferences about objective and subjective risk.

For instance, a participant with belief p that the event will occur believes that lottery L_t occurs with probability p and L_f occurs with probability $1-p$. A participant who treats lotteries involving subjective contingencies just like objective lotteries where the subjective events are replaced with their relevant subjective probability and also has preferences consistent with the reduction of compound lotteries is willing

to trade off between p_t and p_f at a rate of $-\frac{p}{1-p}$: the odds-ratio associated with their belief. Thus, in the case of a belief $p = 0.5$, the indifference curves for (p_t, p_f) are lines with a slope of -1 since an increase to the probability of event a in L_t can be offset by the same reduction in L_f . Under the additional assumption of monotonicity, the optimal choice from the budget set is the bundle $(0.75, 0.75)$. This is, of course, consistent with “truth-telling” in the BQSR (revealing a belief of 0.5).

This demonstrates the delicate nature of the BQSR. The quadratic payoff function ensures that the trade-off between p_t and p_f is equal to the odds-ratio, which is also the rate participants who reduce compound lotteries are willing to trade off between these probabilities.

However, when reduction fails, the indifference curves are no longer linear, breaking incentive compatibility. But this is true only globally— that is for all beliefs $p \in [0, 1]$. Locally, the BQSR might be incentive compatible even when reduction fails. This is particularly true at $p = 0.5$. Suppose we relax reduction and instead assume that indifference curves over (p_t, p_f) are *convex* and *symmetric*. The optimal choice remains $(0.75, 0.75)$ associated with revealing belief $\tilde{p} = 0.5$. We will prove this formally below, but the geometry of the situation is rather intuitive. A convex indifference curve that is symmetric about the 45-degree line is tangent to the concave and symmetric budget only at 0.5.

Convexity ensures that the participant prefers less extreme compound lotteries with respect to the probability of the outcome a in the two sublotteries. Symmetry of the indifference curves requires that $(\pi_1, \pi_2) \sim (\pi_2, \pi_1)$. This is a reasonable assumption *only* in the case of subjective belief $p = 0.5$. This is because the compound lotteries associated with stated beliefs \tilde{p} and $1 - \tilde{p}$ are the same, only changing whether the sublottery with a higher chance of a occurs when the event is true or when it is false. When the participant believes these events have the same probability, it is reasonable to assume the participant has no preference over *when* the more favorable sublottery occurs— since it occurs with the same probability.

We now prove under the assumptions that preferences over (p_t, p_f) are *symmetric*, *monotone*, and *convex*, the optimal choice is $\tilde{p} = 0.5$ or the bundle $(0.75, 0.75)$.

Proof: Consider any other choice of $p_t \in [0, 1] / 0.75$. Assume for contradiction that it is optimal. Since the budget set is $p_f = p_t + 2\sqrt{1 - p_t} - 1$, this bundle is $(p_t, p_t + 2\sqrt{1 - p_t} - 1)$. Since the budget is symmetric about the 45-degree line, the bundle $(p_t + 2\sqrt{1 - p_t} - 1, p_t)$ is also in the budget set. By symmetry of preferences,

$(p_t, p_t + 2\sqrt{1-p_t} - 1) \sim (p_t + 2\sqrt{1-p_t} - 1, p_t)$. By convexity of preferences, the convex combination of these two bundles $(p_t + \sqrt{1-p_t} - \frac{1}{2}, p_t + \sqrt{1-p_t} - \frac{1}{2})$ is preferred to either endpoint. Since $p_t + \sqrt{1-p_t} - \frac{1}{2} < 0.75$ for any $p_t \in [0, 1] / 0.75$, by monotonicity, $(0.75, 0.75) \succ (p_t + \sqrt{1-p_t} - \frac{1}{2}, p_t + \sqrt{1-p_t} - \frac{1}{2})$. By transitivity, $(0.75, 0.75) \succ (p_t, p_t + 2\sqrt{1-p_t} - 1)$ contradicting that the chosen bundle is optimal.

Appendix B: Additional table

	(1)	(2)	(3)
ROE	-0.035 (0.024)		-0.047 (0.033)
Non-reducer		-0.139*** (0.042)	-0.140*** (0.041)
Non-reducer \times ROE			0.024 (0.048)
Stratum FE	Yes	Yes	Yes
Observations	1,206	603	1,206
Adjusted R ²	0.001	0.018	0.017

Table B.1: Accuracy for reducers and non-reducers in the BQSR and ROE tasks, stratum FE specifications

Note: Estimated coefficients are from linear regressions of the accuracy dummy variable on the indicated regressors and stratum fixed effects. Stratum FE correspond to the block-randomization of tasks and task ordering faced by participants. Unit of observation is the participation-elicitation pair. Sample in columns 1 and 3 is all non-centered elicitations. Sample in column 2 is non-centered BQSR elicitations only. Standard errors clustered at the participant level in columns 1 and 3 in parentheses. Heteroskedasticity-robust standard errors in column 2 in parentheses. ***: 0.01, **: 0.05, *: 0.1.

Online Appendix: Experiment Screenshots

Welcome

Please enter your Prolific ID:

Next

Introduction

Thank you for participating!

Please click on the audio file to listen to the instructions. You can follow along in the text below.

Note that the "Next" button on pages with instructions (like this one) will only become available after the audio file has played through. If you prefer to read the instructions, please press the  button to mute the audio while it plays or silence your device.

Press Play



You will complete 6 tasks in this survey. The "Progress Bar" at the bottom of each page tracks how far along you are.

Confidentiality

Your participation in this survey is completely voluntary and your responses are confidential. Your name will not be published and only the research team will use the information collected. There is no risk associated with your participation in this survey. You are free to withdraw from the survey at any time with no penalty beyond the loss of potential earnings.

This survey has been reviewed by the Institutional Review Board (IRB) at Vanderbilt University to ensure your privacy is protected. You may direct any questions or concerns about this survey to the Vanderbilt IRB at (866) 224-8273. For more information, [click here](#).

Payment

For simply completing this 15 to 20 minute survey, you will earn \$3. You may earn an additional \$5 based on the decisions you make and random chance.

At the end of the survey, one of the 6 tasks you complete will be randomly selected by the computer to determine whether you earn the \$5 in addition to the \$3 you earn for completing the survey. Each task has an equal chance of being selected. Your decisions in one task will not affect the potential payment in another task, nor the probability that a task is chosen. You will not know which task is chosen until the end of the survey, so treat each task as if it is the one that determines your payment.

Next

8%

Instructions

Press Play



You will now complete two tasks in which you choose between two payment options. If one of these tasks is selected for payment, the computer will then randomly select two numbers to determine your payment. *Random Number A* determines your chance of earning the \$5. It will be a whole number from 1 to 6 (each number is equally likely to be chosen). *Random Number B* will be a whole number from 1 to 100 (each number is equally likely to be chosen). If *Random Number B* is equal to or less than your chance of earning the \$5, which was determined by *Random Number A*, you earn the \$5.

Here is an example of two payment options:

Option 1

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	50%	50%	50%	50%	50%	50%

Option 2

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	40%	40%	60%	60%	60%	60%

If you chose Option 1, you would have a 50% chance of earning the \$5 regardless of which *Random Number A* (from 1 to 6) is selected. So, you would earn the \$5 if the *Random Number B* (from 1 to 100) was 50 or less. Since the computer is equally likely to pick any number from 1 to 100, this procedure provides exactly the chance of earning the \$5 described by the payment option (in this case, 50%).

If you chose Option 2, you would have a 40% chance of earning the \$5 if *Random Number A* was a 1 or a 2 and a 60% chance of earning the \$5 if *Random Number A* was a 3, 4, 5, or 6. If you had a 40% chance of earning the \$5, you would earn the \$5 if *Random Number B* was 40 or less. If you had a 60% chance of earning the \$5, you would earn the \$5 if *Random Number B* was 60 or less.

Next

17%

Task 1

Remember, if one of these tasks is selected for payment, the computer will then randomly select two numbers to determine your payment. *Random Number A* will be a whole number from 1 to 6 (each number is equally likely to be chosen) that determines your chance of earning the \$5. *Random Number B* will be a whole number from 1 to 100 (each number is equally likely to be chosen). If *Random Number B* is equal to or less than your chance of earning the \$5, determined by *Random Number A*, you earn the \$5.

Select the lottery you prefer to determine your payment.

Option 1

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	75%	75%	75%	75%	75%	75%

Option 2

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	55%	55%	89%	89%	89%	89%

Next

25%

Task 2

Remember, if one of these tasks is selected for payment, the computer will then randomly select two numbers to determine your payment. *Random Number A* will be a whole number from 1 to 6 (each number is equally likely to be chosen) that determines your chance of earning the \$5. *Random Number B* will be a whole number from 1 to 100 (each number is equally likely to be chosen). If *Random Number B* is equal to or less than your chance of earning the \$5, determined by *Random Number A*, you earn the \$5.

Select the lottery you prefer to determine your payment.

Option 1

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	75%	75%	75%	75%	75%	75%

Option 2

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	31%	97%	97%	97%	97%	97%

Next

33%

Instructions

Press Play



In the next two tasks, you will be asked to rank one payment option relative to other payment options in order from your most preferred payment option to your least preferred payment option. The other payment options will be pre-ordered from highest likelihood of winning \$5 to lowest likelihood of winning \$5. We'll call this list of pre-ordered payment options the "Pre-ordered List." The one payment option you must choose how to rank is called the "Unordered Option."

If this task is selected for payment, one of the options from the Pre-ordered List will be randomly selected by the computer. Whichever you rank higher, the randomly selected option from the Pre-ordered List or the Unordered option, will determine your payment.

For example, consider an urn filled with four balls. Three of those balls are red and one of those balls is green. You could be asked to rank the following Unordered Option:

Unordered Option

- You are paid \$5 if a randomly selected ball is red.

...in the following Pre-ordered List:

Pre-ordered List

- You are paid \$5 with 80% chance.
- You are paid \$5 with 50% chance.
- You are paid \$5 with 10% chance.

You need to choose where to rank the Unordered Option relative to the payment options in the Pre-ordered List. For example, you might choose the following ranking:

1. You are paid \$5 with 80% chance.
2. You are paid \$5 if a randomly selected ball is red.
3. You are paid \$5 with 50% chance.
4. You are paid \$5 with 10% chance.

If one of the options from the Pre-ordered List determines your payment, based on your rankings and the computer's random selection, the computer will randomly select a whole number from 1 to 100 (each number is equally likely to be chosen). If that number is equal to or less than your chance of being paid the \$5, you earn the \$5. Otherwise, you do not earn the \$5. Since the computer is equally likely to pick any number from 1 to 100, this procedure provides exactly the chance of being paid \$5 described by the option.

If the Unordered Option determines your payment, based on your rankings and the computer's random selection, the computer will execute the action described in a random and fair manner. You earn the \$5 if the event occurs as described in the option. In the above example, if you were to be paid based on the option "You are paid \$5 if a randomly selected ball is red," the computer would randomly select a ball from the urn filled with three red balls and one green ball (each ball is equally likely to be chosen). If the ball was red, you would earn the \$5. Otherwise, you would not earn the \$5.

Suppose you ranked "You are paid \$5 if a randomly selected ball is red" **below** "You are paid \$5 with 50% chance," instead of how these two payment options are ranked in the example. Then you would have a lower chance of earning the \$5 if "You are paid \$5 with 50% chance." is the randomly selected option from the Pre-Ordered List. In this case, you would be paid based on a 50% chance because you ranked it higher, whereas there is a 75% chance of earning the \$5 if you are paid based on a selected ball being red (3 balls out of 4). So, **you are most likely to earn the \$5 if you rank the payment options from highest likelihood to lowest likelihood of occurring.**

Next

42%

Task 3

Please rank the Unordered Option on the left in the Pre-ordered List on the right by dragging the Unordered Option to a green box.

Unordered Option
You are paid \$5 if a fair, six-sided die rolls a 2, 4, or 6.

Pre-ordered List
<i>Highest-ranked</i>
You are paid \$5 with 100% chance
You are paid \$5 with 95% chance
You are paid \$5 with 90% chance
You are paid \$5 with 85% chance
You are paid \$5 with 80% chance
You are paid \$5 with 75% chance
You are paid \$5 with 70% chance
You are paid \$5 with 65% chance
You are paid \$5 with 60% chance
You are paid \$5 with 55% chance
You are paid \$5 with 50% chance
You are paid \$5 with 45% chance
You are paid \$5 with 40% chance
You are paid \$5 with 35% chance
You are paid \$5 with 30% chance
You are paid \$5 with 25% chance
You are paid \$5 with 20% chance
You are paid \$5 with 15% chance
You are paid \$5 with 10% chance
You are paid \$5 with 5% chance
You are paid \$5 with 0% chance
<i>Lowest-ranked</i>

Task 4

Please rank the Unordered Option on the left in the Pre-ordered List on the right by dragging the Unordered Option to a green box.

Unordered Option
You are paid \$5 if a fair, six-sided die rolls a 5 or 6.

Pre-ordered List
<i>Highest-ranked</i>
You are paid \$5 with 100% chance
You are paid \$5 with 95% chance
You are paid \$5 with 90% chance
You are paid \$5 with 85% chance
You are paid \$5 with 80% chance
You are paid \$5 with 75% chance
You are paid \$5 with 70% chance
You are paid \$5 with 65% chance
You are paid \$5 with 60% chance
You are paid \$5 with 55% chance
You are paid \$5 with 50% chance
You are paid \$5 with 45% chance
You are paid \$5 with 40% chance
You are paid \$5 with 35% chance
You are paid \$5 with 30% chance
You are paid \$5 with 25% chance
You are paid \$5 with 20% chance
You are paid \$5 with 15% chance
You are paid \$5 with 10% chance
You are paid \$5 with 5% chance
You are paid \$5 with 0% chance
<i>Lowest-ranked</i>

Instructions

Press Play



In the next two tasks, you will be asked to guess the chance that some event occurs. Your guess is a percentage probability from 0 to 100, with 0 indicating a 0-out-of-100 chance that the event occurs and 100 indicating a 100-out-of-100 chance that the event occurs. The number you provide is called *Your Guess*.

You choose *Your Guess* by clicking the response bar on your screen. The width of the green part of the bar indicates your guess that the event occurs.

- Larger values of *Your Guess* represent a greater chance that the event occurs and a smaller chance that the event does not occur.
- Smaller values of *Your Guess* represent a smaller chance that the event occurs and a greater chance that the event does not occur.

The gray part of the bar is $100 - \textit{Your Guess}$, and represents your guess that the event does not occur.

Payment Rule

We now explain how *Your Guess* is used to determine whether you earn the \$5. **The payment rule is designed so that you can secure the largest chance of earning the \$5 by reporting your most-accurate guess.**

- The computer chooses two numbers between 1 and 100, where each number is equally likely, as if rolling two 100-sided dice. These numbers are called Computer Number A and Computer Number B.
- The computer determines whether you earn the \$5 according to whether the event occurs.

The event occurs: You will earn the \$5 if *Your Guess* is **greater than or equal** to **either** Computer Number A or Computer Number B or both numbers.

The event does not occur: You will earn the \$5 if *Your Guess* is **less than or equal** to **either** Computer Number A or Computer Number B or both numbers.

To help you understand the payment rule, as you move the slider to choose *Your Guess*, the computer will inform you of:

- The probability of earning the \$5 if the event occurs.
- The probability of earning the \$5 if the event does not occur.

For example, consider an urn filled with four balls. Three of those balls are red and one of those balls is green. You could be asked to guess whether a randomly selected ball from the urn is red. Suppose *Your Guess* was 75% and the ball selected was red. If either of the computer numbers is 75 or *lower*, then you would earn the \$5. If both are *above* 75, you would not earn the \$5. On the other hand, suppose your guess was 75% and the ball selected was *not* red. If either of the computer numbers is 75 or *higher*, then you would earn the \$5. If both are *below* 75, you would not earn the \$5.

Next

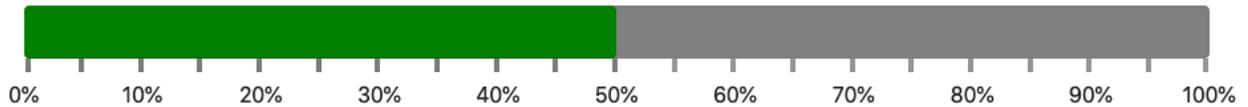
67%

Task 5

If this task is selected for payment, the computer will roll a fair six-sided die. Each number is equally likely to be rolled.

What is *Your Guess* about the chance that the die rolls a 1, 3, or 5?

Current Guess: 50%



Your probability of earning the \$5 if the die roll is a 1, 3, or 5: 75%

Your probability of earning the \$5 if the die roll is a 2, 4, or 6: 75%

Next

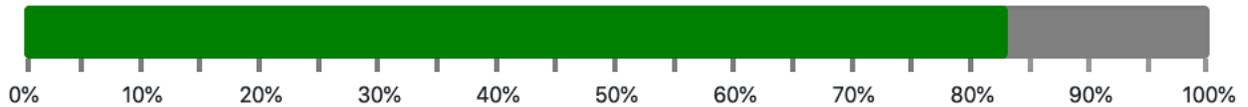
75%

Task 6

If this task is selected for payment, the computer will roll a fair six-sided die. Each number is equally likely to be rolled.

What is *Your Guess* about the chance that the die rolls a 1, 2, 3, 4, or 5?

Current Guess: 83%



Your probability of earning the \$5 if the die roll is a 1, 2, 3, 4, or 5: 97%

Your probability of earning the \$5 if the die roll is a 6: 31%

Next

83%

Payment

Thank you for participating! You have earned \$3 for completing the survey. Now, click the button below that says "Choose Task" to determine which of the 6 tasks is randomly selected by the computer to determine your payment today.

Choose Task

92%

Payment

Task 1 has been selected for payment.

In Task 1, you indicated you prefer the following payment option:

If <i>Random Number A</i> is:	1	2	3	4	5	6
Your chance of earning the \$5 is:	75%	75%	75%	75%	75%	75%

Random Number A is 4, so your chance of winning \$5 is 75%. Random Number B will be a number between 1 and 100. You will win \$5 if Random Number B is 75 or less. Click the button below to randomly determine Random Number B.

Random Number B is 87, so you do not win \$5. You earned \$3 simply for participating, so your total payment is \$3.

Click "Next" for instructions on how to record your payment.

Next

92%

Please click the link below to finish the survey.

<https://app.prolific.co/submissions/complete?cc=107C59DA>

You will be redirected to the Prolific website and your payment will be recorded.

100%