



Perceptions, Intentions, and Cheating

Li Hao and Daniel Houser

February 2013

Discussion Paper

Perceptions, Intentions, and Cheating

Li Hao and Daniel Houser^{*}

February 18, 2013

Abstract: We report data from a laboratory experiment demonstrating that cheating is significantly deterred when a possible intent to cheat must be revealed *before*, rather than *after*, a potentially dishonest act. Further, data from independent evaluators suggests a reason: the same action is more likely to be perceived as dishonest when cheating could have been planned, as compared to cases when it seems simply impulsive. Overall, we find the vast majority of participants prefer to appear honest, but only a minority prefers actually to be honest. Finally, we conduct a type-classification analysis that implies that after establishing an honest appearance people cheat to the greatest extent possible. It follows that the “incomplete cheating” behavior frequently reported in the literature may be due more to a preference for maintaining an honest appearance than an intrinsic aversion to cheating.

JEL: C91; D03

Keywords: cheating; honest appearance; partial cheating; experimental design

^{*} Li Hao: email: lhao@walton.uark.edu; office: (+1) 479 575 6187; fax: (+1) 479 575 3241. Department of Economics, Walton College of Business, WCOB 421, University of Arkansas, Fayetteville, AR 72701
Daniel Houser: email: dhouser@gmu.edu; office: (+1) 703 993 4856; fax: (+1) 703 993 4851. Interdisciplinary Center for Economic Science, George Mason University, 4400 University Dr, MSN 1B2, Fairfax, VA 22030.

Some of the data and analyses reported in this paper were previously reported in the unpublished working paper “Honest Lies,” by the same authors. For helpful comments we thank Glenn Harrison, Edi Karni, R. Lynn Hannan, Cary Deck, Roberto Weber, Omar Al-Ubaydli, Marco Castillo, Ragan Petrie, Jo Winter, Larry White, Chris Coyne, our colleagues at ICES, George Mason University, seminar participants at CEAR, Georgia State University (2010), the ESA North-American meeting (2010), University of Fayetteville, Arkansas (2011), and GSPW at George Mason University (2011). The authors are of course responsible for any errors in this paper.

“There is a return to *appearing* honest, but not to *being* honest.”
- Akerlof (1983, p. 57)

1. Introduction

People prefer to appear honest, perhaps especially those who occasionally cheat. The reason, as observed by Akerlof (1983), is that people benefit from appearing honest, but do not benefit from actually *being* honest. Indeed, people are willing to pay to maintain an honest appearance, as evidenced by the market for alibis and excuses for absences¹. Despite the importance people attach to appearing honest, as well as a substantial recent economics literature on deception (e.g., Gneezy (2005) and papers reviewed in section II below), we are aware of no published research that studies the interactions between preferences for appearing and being honest².

To identify separate preferences for *appearing* and *being* honest, the key is the revelation of intent to cheat. By definition, the appearance of honesty is immediately jeopardized when the intent to cheat is revealed. The attribution of intent is critical in dealing with deception, fraud and crime, a fact which has long been established in the psychology literature (Schweitzer et al. 2006; Stouten et al. 2006). Deceivers with no evidence of having cheated *intentionally* are more likely to be forgiven, and consequently avoid retribution (Hippel and Trivers 2011). A familiar example from criminal law is that the intent to cause bodily harm, and potentially death, is required for a homicide to be murder (p. 85, Mewett, 1978).

Moreover, the intent can be developed well beforehand, or at the moment of the action. This time element, as it turns out, makes a critical difference. For example, first-degree murderers face much heavier charges and punishments than second-degree murderers; what elevates a murder to first-degree is whether it was *planned* beforehand and considered *not impulsive* on the part of the accused (Felson and Massoglia 2011). Naturally, planning and deliberating requires

¹ For example, the company “Alibi Network”(www.alibinetwork.com) offers customized alibis to clients. The company provides fabricated airline confirmation, hotel stay and car rental receipts for any location and time of the client’s choice. For those who want excuses for an upcoming absence, a 2-5 day alibi package is offered so that one can pretend he/she is going to a conference or career training. The package is extremely comprehensive and individually tailored, including the conference invitation, confirmation emails and/or phone calls, mailed conference programs such as timetable and topic overview, virtual air ticket and hotel stay confirmation, and even a fake hotel number that is answered by a trained receptionist.

² Jiang (2012) also attempts to separate preferences for appearing and being honest, and to do so uses a strategy similar to what we report in this paper. The two studies were developed independently. We review Jiang’s (2012) study in Section 2.

thought and time. The evidence that one has spent time on premeditation and consideration is critical when assessing their misconduct. It is this time element that we use as the main treatment manipulation in our study, as we describe below.

Our investigation is based on a novel game that requires participants to predict die roll outcomes. We consider three treatments: “Control,” “Impulsive,” and “Planned.” In Control, subjects first predicted the probability that each outcome would occur, and submitted the predictions on paper to the experimenter. Then, they each rolled a die, and the experimenter observed and recorded the outcomes. Subjects’ predictions were evaluated according to their own die roll outcomes, with more accurate predictions earning more money. There was no opportunity to cheat for profit.

In contrast, the Impulsive and Planned treatments allowed subjects to increase their earnings by cheating. The Impulsive treatment differed from Control in that subjects did not submit their predictions until after they rolled the die. Hence, they could maximize profits by discarding their original predictions and instead reporting that they predicted that the observed outcome would occur with 100% probability. On the other hand, the Planned treatment differed from Control in that subjects rolled the die *privately*. In particular, subjects were assured before submitting the prediction that they would later roll the die privately and self-report the outcome to the experimenter³. Here, subjects could maximize profits by planning ahead, *i.e.*, first submitting a prediction that a particular outcome would occur with 100% probability, and then reporting the matching outcome regardless of the actual outcome.

The profit maximizing strategies in Impulsive and Planned were identical: “correctly” predicting the actual die roll outcome occurs with certainty, by either misreporting the original prediction or the actual die roll outcome. However, a key difference between the two treatments is the time element in revealing the intent to cheat. In Planned, the cheating action was misreporting the private die roll outcome, and the intent to do so was embodied in any prediction that differed significantly from the objective uniform distribution over possible outcomes. Because subjects’ predictions were submitted well before self-reported die roll outcomes, any

³ Due to its sensitive nature, the possibility of cheating was not explicitly announced. However, the fact that the die roll would be private was emphasized three times in the instructions. Hard copy instructions were in front of the subjects during the entire experiment, and were also read aloud by the experimenter.

dishonest-looking predictions could be perceived as a signal that the participant spent time planning to cheat. In contrast, in Impulsive, the cheating action occurred via misreporting predictions, and subjects submitted their predictions immediately after their die roll outcome. Consequently, there was no clear “evidence” that any cheating was premeditated.

Our key question is whether people are less willing to cheat when doing so negatively impacts the appearance that they are honest. We investigate this by comparing behaviors between the Planned and Impulsive treatments. As we expect cheating in Planned to impact appearances more negatively, we expect less cheating to occur in Planned than in Impulsive.

To provide evidence that cheating is in fact perceived differently between the Planned and Impulsive treatments, we collected opinions from independent evaluators regarding the appearance of honesty in Impulsive and Planned. In particular, participants were asked to evaluate identical scenarios from Impulsive and Planned, and were rewarded if their answers matched the majority’s answers. Our key finding is that *ceteris paribus*, the same action is more likely to look like cheating in Planned than Impulsive. This suggests the perception of honesty is dampened by evidence that cheating could have been premeditated, a possibility which we manipulate by the timing of the revelation of intention. It is perhaps worth noting that, to our knowledge, the use of external evaluators to assess the appearance of honesty is novel to our study (e.g., this approach was not followed in the closely related paper by Jiang, 2012).

We find less cheating in Planned than Impulsive. To shed further light on this result, we conducted a type-classification analysis using an algorithm (adapted from El-Gamal and Grether, 1995), and found the vast majority of subjects (95%) exhibited a strong preference for appearing honest, while at the same time less than half of subjects were truly honest. The prevalent preference for appearing honest offers important insights for designing institutions to deter misconduct, especially when monitoring or contracting on all possible contingencies is too costly or even infeasible (Williamson 1975). The key idea is to provide incentives for people to reveal their plans *before* actions, making it harder to hide the intent to cheat. For instance, employees should be rewarded for submitting travel budgets before their business trips, or managers rewarded for submitting a cost budget before starting new projects. Similarly, government administrations can gain more voters’ support by increasing transparency, which in turn may mitigate corruption. This idea may also be a step towards addressing self-control problems (that

is, cheating on oneself). For instance, when trying to lose weight it may be helpful to share diet and exercise plans and offer regular progress reports to trusted others.

Our results help to explain the puzzle of incomplete cheating. Previous research strongly suggests that people of all ages are averse to lying (see, *e.g.*, Gneezy 2005; Hannah et al. 2006; Bucciol and Piovesan 2008; Fischbacher and Heusi 2008; Greene and Paxton 2009; Mazar et al. 2008; Houser et al. 2010; Lundquist et al. 2009). One interesting and persistent pattern reported in these studies is that, when given the opportunity, people do cheat but shy away from cheating for maximum earnings. Our results suggest that a source of incomplete cheating may be more tied to a desire to maintain an honest appearance, rather than an intrinsic aversion to cheating.

The paper proceeds as follows. Section 2 reviews related literature; Section 3 describes the design of the experiment; Section 4 specifies our hypotheses; Section 5 reports the results; and the final section concludes the paper.

2. Literature Review

In recent years, there has been a surge of research in cheating behavior. In this section, we briefly review the papers most relevant to our study.

Gneezy (2005) showed that people exhibit an aversion to lying. In his sender-receiver game, only Player 1 was informed about the monetary consequences of the two options, and Player 2 chose which option should be implemented based on the message sent from Player 1. Hence, Player 1 could either: (i) tell the truth and obtain Option A, in which his payoff was lower than Player 2's; or (ii) lie and obtain Option B for a slight monetary gain at a greater cost to Player 2. In an otherwise identical dictator game, Player 1 chose between Options A and B; Player 2 had no choice but to accept the payoff division. The paper reported that the proportion of Option B was significantly lower in the sender-receiver game than in the dictator game, thus suggesting an aversion to lying as opposed to preferences over monetary allocations. In addition, Gneezy (2005) also found that people lie less when the lie results in a greater cost to others.

Gneezy's (2005) findings stimulated subsequent work that reported consistent results (see, *e.g.*, Sánchez-Pagés and Vorsatz 2007, 2009; Lundquist et al. 2009; Rode 2010; Hurkens and

Kartik (forthcoming)). For example, Lundquist et al. (2009) found that lying aversion is greater when the size of the lie (*i.e.*, the difference between the truth and the lie) is greater. In their experiment, Player 1 reported his type to Player 2, who decided whether to enter into a contract with Player 1. Upon completing the contract, Player 1 always gained. Player 2 gained if Player 1's type was above a threshold, but otherwise lost. The authors found that the further Player 1's type was from the threshold, the less likely he would lie about his type.

Mazar et al. (2008) argue a theory of self-concept maintenance; they observe that “people behave dishonestly enough to profit, but honestly enough to delude themselves of their own integrity.” The authors suggest two mechanisms that allow for such self-concept maintenance: (i) inattention to moral standards; and (ii) categorization malleability. For example, in one of their experiments, subjects self-reported their own performance on a real-effort task, and were paid accordingly. However, some subjects were asked to write down the Ten Commandments before the task, while others were not. The result is that those who were reminded of moral standards lied less, thus supporting the hypothesis that inattention to moral standards serves as a mechanism through which people cheat for profit without spoiling a positive self-concept.

In Fischbacher and Heusi's (2008) experiment, subjects rolled a six-sided die privately and self-reported the first roll. The outcome of the first roll was the amount of payment they received for the experiment. The fraction of self-reported highest payoff outcomes was significantly higher than one sixth, as expected; however, the fraction of the second highest payoff was also significantly higher than one sixth. This is a type of “incomplete cheating,” which the authors speculate might be due to greed aversion and the desire to appear honest.

Building on Fischbacher and Heusi (2008), Jiang's (2012) experiment shares a similar idea to ours, although developed independently. In Jiang's “mind game” experiment, subjects rolled a six-sided die, and denoted the outcome by x , where $x=1, 2, 3, 4, 5$, or 6 . Subjects also chose between two earning schemes, so that their earnings were either (i) x euros or (ii) $6-x$ euros. There were two treatments. In the “report-first” treatment, subjects chose the earning scheme before the die roll; in the “throw-first” treatment, subjects observed the die roll outcome before choosing the earning scheme. Hence, “report-first” and “throw-first” are similar to our Planned and Impulsive treatments, respectively. Furthermore, Jiang (2012) found that subjects cheated more in “throw-first” than in “report-first”, which is consistent with our results. Our experiment

differs from Jiang (2012) in three ways. First, our participants were paid based on a QSR, while theirs were paid based only on die roll outcomes. Second, we provide a baseline, the Control treatment, where subjects cannot cheat. Finally, the independent evaluations of Planned and Impulsive provide evidence on the validity of our manipulation.

3. Experimental Design

Our key innovation lies in the subtle difference between Planned and Impulsive. In Planned, the intent to cheat was manifested as predictions that differ drastically from the objective distribution (*i.e.*, uniform). Even worse, predictions were submitted well before the action of cheating on private die roll outcomes, which sent a strong signal that the cheating was premeditated. In Impulsive, however, predictions were submitted after the die roll outcome, and cheating occurred via mis-reporting predictions. Hence, there was no obvious evidence that cheating was pre-conceived in a way that was planned beforehand.

Another novelty of our experiment is the use of independent evaluators who were incentivized to evaluate the cheating behavior in the main treatments. By comparing these evaluations and the cheater's behavior in Planned and Impulsive, we gain a deeper understanding about the cheating behavior and perceptions.

Subjects were recruited via email from registered students at George Mason University. The experiment was conducted between June 2009 and March 2012. Upon arrival, subjects were seated in individual cubicles, separated by partitions, so that others could not observe their actions. Sessions lasted 40 minutes on average, and earnings ranged between \$6.25 and \$25, in addition to a show-up bonus of \$5.

3.1. Design

Our main experiment included three treatments: Control, Impulsive, and Planned. After the treatments, a final group of subjects were asked to evaluate whether participants in given

scenarios from Impulsive and Planned treatments appeared to have cheated. We used a between-subject design, so no one participated in more than one session from the main treatments and the evaluation sessions. All instructions and decisions sheets are attached in Appendix 1.

Control treatment. There were two stages. In the first stage, the subject predicted the percent chance that each face of a fair four-sided die would turn up in a single roll, and then submitted his/her prediction to the experimenter on paper. The individual probabilities had to be between 0% and 100% (inclusive), and the four probabilities had to add up to exactly 100%.

In the second stage, the subject rolled a fair four-sided die. The experimenter observed and recorded the outcome. Subjects earned more for accurate prediction, according to a quadratic scoring rule detailed in the next subsection.

Impulsive treatment. Identical to Control, except that in the first stage, the subject was asked to keep his predictions in mind, and wait until *after* the die roll outcome was observed and recorded in the second stage. Then, the subject recalled his predictions and submitted them on paper.

Planned treatment. Identical to Control, except that the die roll in the second stage became private, and the subject knew before he made predictions that his rolling would not be monitored by the experimenter. To make subjects comfortable, they were encouraged to roll the die privately as many times as they wished, but to only remember the first roll and report it to the experimenter via paper.

A few features of our design are worth noting. First, we did not explicitly invite subjects to cheat. We did, however, fully endeavor to ensure that subjects understood they had opportunity to cheat. This is especially important in the Planned treatment (while at the same time doing our best to avoid any experimenter demand effects), as it was emphasized three times that the die roll would be private—at the beginning, the middle, and the end of the instructions. Hard copy instructions were at the subjects' disposal during the entire experiment. Additionally, the experimenter illustrated the experiment with five examples, including the earnings-maximizing strategy. A comprehension quiz was conducted, and subjects were required to answer all questions correctly before proceeding to the experiment.

We also announced to subjects that the dice were fair; thus, predictions that differed from the objective distribution could not be attributed to suspicions that the dice might be biased. However, we also encouraged subjects to play hunches if they believed certain outcomes were more likely than others. The goal was to ensure subjects felt comfortable making predictions other than the uniform probability prediction⁴.

Evaluation sessions. A final group of subjects served as independent observers to evaluate whether a participant in a given scenario from the main treatments appeared honest or not. These evaluators received detailed debriefing about the Impulsive and Planned treatments, including the original instructions (for details see Appendix 1).

There were two parts to their evaluation task. The first part consisted of evaluating twelve scenarios, where highest-payoff outcomes were obtained in all cases. Scenarios 1 to 6 were all from the Planned treatment, and differed only according to p_i^{max} values, which were 50%, 60%, 40%, 25%, 85%, and 100%, respectively⁵. Scenarios 7 to 12 were from the Impulsive treatment, repeating the same sequence of p_i^{max} as the first six scenarios. To avoid any framing effects, we used “Experiment 1” and “Experiment 2” for the “Planned” and “Impulsive” treatments, respectively. For example, the first scenario was as below,

Scenario 1: In Experiment 1, a participant first revealed his or her prediction that outcome 2 would occur with 50% chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die privately, and reported that the die roll was 2.

*Question 1: Do you think the majority of people in this room believe that this participant cheated?
 (“Yes” or “No”)*

The second part included 6 questions asking evaluators to directly compare identical scenarios between the two treatments, e.g., “If your answers in **Scenarios 1 and 7** are the same, do you think the majority of people in this room believe that the participant more likely to have

⁴ Reasons to regret the uniform prediction include the desire not to appear “ignorant,” as well as potential experimenter demand effects.

⁵ We first chose 6 key p_i^{max} that spanned from its minimum value of 25% to its maximum 100%. We then randomized the order in which they were presented to subjects. Note that $p_i^{max} = 25\%$ indicates objective predictions, and evaluators should not have viewed this prediction as dishonest. We included this one to verify that subjects understood the experiment.

cheated in Scenario 1 or in Scenario 7?” Subjects could choose among three answers: “Scenarios 1”, “Scenarios 7,” and “equally likely.”

Three questions, one from each set of six questions, were randomly selected for payment, and subjects earned \$5 for each answer that was in agreement with the majority of the evaluators in the room.

3.2. Payoff in the Main Experiment

In the main experiment, each subject’s first-stage probabilistic prediction was compared with the relevant roll in the second stage. Earnings were calculated according to the following quadratic scoring rule, which rewards prediction accuracy.

$$Earnings = \$25 - \$12.5 \sum_{i=1}^4 (\chi_i - p_i)^2 \quad (1)$$

where i indexes the four faces of the die: $i \in \{1, 2, 3, 4\}$, and p_i is the probability that the subject assigned to face i . The indicator χ_i is 1 if face i is the outcome of the roll, and 0 otherwise (so we have $\sum_{i=1}^4 \chi_i = 1$). In the first stage, the subject submitted a vector of four probabilities: $p = (p_1, p_2, p_3, p_4)$, where $0 \leq p_i \leq 1$, and $\sum_{i=1}^4 p_i = 1$.

Quadratic scoring rules are widely used incentive-compatible mechanisms for eliciting subjective probabilities in experimental studies (see *e.g.*, Nyarko and Schotter 2002; Andersen et al. 2010). Kadane and Winkler (1988, p. 359) showed that an expected utility maximizer would report truthfully, assuming the individual’s utility is linear in money⁶.

To facilitate subjects’ understanding of the payoffs, we provided an interactive Excel tool in which subjects could type in any probabilistic prediction and view the payoffs conditional on the rolling outcome (a screenshot is reproduced in Appendix 2)⁷.

4. Hypotheses

We start with a few definitions. A prediction is an “**objective prediction**” if it is identical to the objective distribution (25%, 25%, 25%, 25%). For subject i ’s prediction of four probabilities,

⁶ The other assumption, the no-stakes condition, is not violated here, because subjects’ wealth outside the laboratory experiment is independent of the outcome of the die roll.

⁷ We thank Zachary Grossman for providing us the original version of this tool.

we sort them from highest to the lowest and denote the **highest probability** by p_i^{max} , and the corresponding outcome that the subject assigned p_i^{max} to is called a “**highest-payoff outcome**”⁸. The highest probability and its corresponding outcomes are key variables, as the only way to increase earnings is to concentrate more probabilities on fewer outcomes.

We use predictions in Control to define what is considered “honest-looking.” For subjects who desire to appear honest, a simple strategy is to make predictions as if they are not being monitored. Intuitively, a prediction in Planned would be “honest-looking” if it did not differ from “typical” predictions in Control. To define “typical,” we must draw inferences from the empirical predictions in Control. The highest probability p_i^{max} is critical for a prediction to appear honest because it determines the highest possible payoff. We examine the distribution of p_i^{max} , and find that 99th percentile is 50%; thus, the vast majority of subjects (69 out of 70) in Control stated 50% or less as the highest probability p_i^{max} . We therefore use 50% as the upper bound for “typical” or “honest-looking” predictions⁹. Hence, we call a prediction “**honest-looking**” if it assigns no more than 50% probability to any single outcome of the die roll.

Our first hypothesis concerns the preference for appearing honest, which is made more salient in Planned than in the Impulsive treatment. We predict that the Planned treatment has at least as many honest-looking predictions as the Impulsive treatment.

Hypothesis 1 (Preference for Appearing Honest). *The fraction of honest-looking predictions in Planned is greater than or equal to that in the Impulsive treatment.*

Next, we predict that people cheat when they are given opportunities to do so, as in our Impulsive and Planned treatments. Hence, our second hypothesis follows:

Hypothesis 2 (Preference for Being Honest). *Cheating occurs when given the opportunity.*

⁸ In the event of ties, there are multiple highest-payoff outcomes.

⁹ This threshold says that roughly 99% of the time, a random draw from Control is no greater than 50%. In the Control treatment, the highest prediction is 57%, which is followed by two predictions at 50%, and quite a few predictions between 50% and 45%. Hence, 50% seems a natural focal point that subjects in Control were comfortable with. The 95th percentile of the distribution of p_i^{max} in Control is 48%. If we instead use the 95th percentile to define the upper bound of “typical” or “honest-looking” predictions, 90% of predictions in *Planned* were “honest-looking.”

Finally, we compare the level of cheating between Impulsive and Planned, and predict that cheating is deterred in Planned compared to the Impulsive treatment. The reason is that individuals in Impulsive can more easily hide evidence that their cheating was premeditated. Hypothesis 3 is summarized as below,

Hypothesis 3 (Preference for Being Honest). *The level of cheating in Impulsive is greater than or equal to cheating in Planned.*

5. Results

A total of 255 subjects participated in our experiment, including 48 independent evaluators, and 70, 61 and 76 subjects in Control, Impulsive, and Planned, respectively. Due to the nature of the experiment, we do not, and will never, know whether a particular individual cheated. However, we are able to make inferences about whether subjects in a treatment cheated at all by comparing their die roll outcomes to the objective distribution from rolling a fair die. We are also able to infer whether subjects in one treatment cheated more than those in another treatment.

Our results are organized as follows. First, we analyze independent evaluations, and find that the same behavior is perceived as less honest in Planned than in Impulsive.

Second, we study the preference for *appearing* honest by comparing subjects' probability predictions in Planned and Impulsive against an "honest-looking" prediction defined by predictions in Control.

Third, we measure the preference for *being* honest via the accuracy of probability predictions and subjects' earnings. Given the opportunity to cheat, if subjects' predictions about die rolls were significantly more "accurate" than those in Control (which implies higher earnings), we can conclude with confidence that cheating occurred.

Finally, we estimate the types of honesty using a variation of El-Gamel and Grether's (1995) algorithm.

5.1. Independent Evaluations of Honest Appearance in Planned and Impulsive

This subsection reports 48 independent evaluators' perceptions of the Decision Maker's (DM hereafter) honesty in the Planned and Impulsive treatments.

We first conduct a consistency check for the evaluations, focusing on two types of significant inconsistencies: (i) stating that objective predictions look dishonest; or (ii) a higher p_i^{max} looks more honest than a lower p_i^{max} . 8 subjects fall into at least one of these two types, and thus are excluded from our analysis¹⁰.

Our first result lends support to the notion (in Section 4) that an "honest-looking" prediction's highest prediction p_i^{max} does not exceed 50%. Figure 1 summarizes evaluations from the first twelve questions regarding whether the DM appears to have cheated, in light of his/her highest prediction p_i^{max} . The horizontal axis shows p_i^{max} in the scenario, and the vertical axis indicates the fraction of evaluators who believe the DM has cheated. We find that once the DM's p_i^{max} surpassed 50%, the majority of the evaluators agreed that the DM appeared to have cheated. Hence, Result 1 follows:

Result 1. (Upper bound for honest-looking predictions). *50% is the cutoff p_i^{max} that the majority of evaluators agree on as the upper bound for honesty-looking predictions.*

Evidence. Figure 1 shows that in both treatments, when p_i^{max} was 40% or below, less than half of evaluators thought the DM had cheated; however, when p_i^{max} was 50% or above, the *majority* of evaluators believed the DM had cheated. Hence, these independent evaluations provide external support to our use of 50% as the maximum prediction the DM can make without seriously damaging his/her honest appearance.

¹⁰ Our results remain largely the same if we do not exclude these 8 inconsistent subjects.

Figure 1: Fraction of evaluators who believe cheating occurred at given p_i^{max}

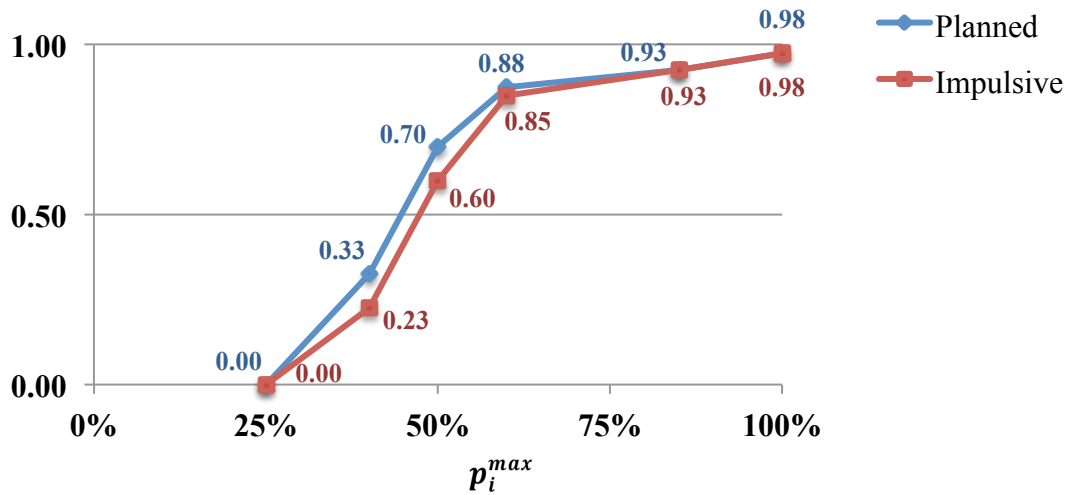
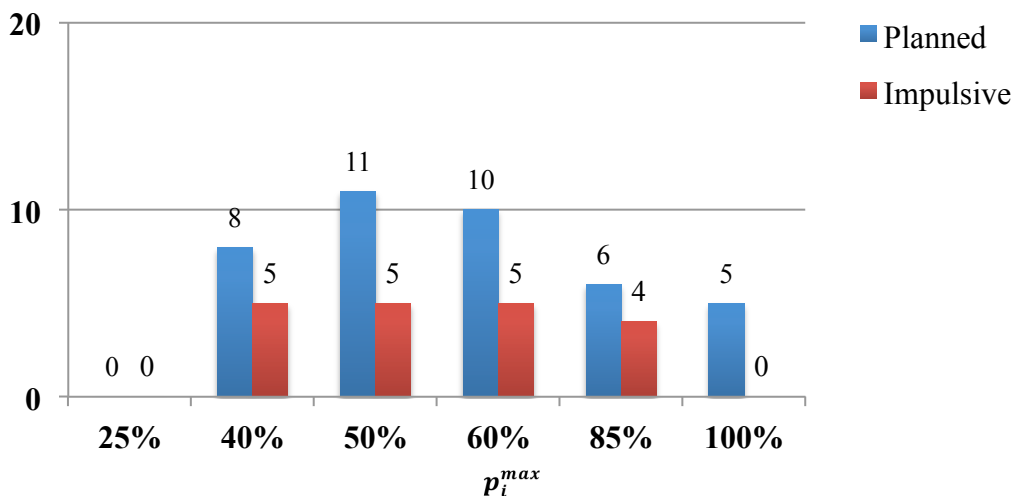


Figure 2 reports data from the final set of six questions where evaluators compare scenarios with identical p_i^{max} between the two treatments, and choose the one that is more likely to be dishonest. The horizontal axis lists all six scenarios by its p_i^{max} . The blue (red) bar indicates the number of evaluators who believed the DM in Planned (Impulsive) was more likely to have cheated than the DM in Impulsive (Planned)¹¹.

Figure 2: Number of evaluators who consider behavior in Planned (Impulsive) more likely to be dishonest than the same behavior in Impulsive (Planned) (N=40)



¹¹ The number of observers who considered the identical actions in the two treatments equally (dis)honest is the difference between 40 and the sum of numbers indicated by the red and blue bars.

Figure 2 shows that, in all scenarios, more evaluators considered actions in Planned more likely to be dishonest than Impulsive. This leads to Result 2.

Result 2. (Independent Evaluations). *Significantly more evaluators considered an identical action in Planned more likely to be dishonest than in Impulsive.*

Evidence. We construct a variable “ChosePlanned”, which is 1 (-1) if the evaluator chooses the Planned (Impulsive) treatment more likely to be dishonest, and 0 if the evaluator is indifferent. Hence, if the mean of ChosePlanned is greater (less) than zero, we can conclude that an identical action in Planned (Impulsive) is more likely to be perceived as dishonest. To account for multiple evaluations from the same evaluator, we run the following GLS regression, clustered by individuals (evaluators).

$$\text{ChosePlanned} = \alpha + \varepsilon, \quad \text{where } \varepsilon \text{ is i.i.d. } \sim N(0,1).$$

The estimated intercept α is the mean of ChosePlanned, after taking into account that we have multiple decisions from the same individual. We find that ChosePlanned is 0.0875 (robust standard error=0.0391), significantly greater than 0 ($p=0.025$, two-sided t-test). Hence, significantly more evaluators considered the Planned treatment to be less honest-looking than Impulsive.

5.2. Preference for Appearing Honest

Based on predictions in the Control treatment, as well as independent evaluations (Result 1), a prediction is considered “honest-looking” if its p_i^{max} does not exceed 50%. Using this definition, we compute the fraction of honest-looking predictions in Planned and Impulsive, and present our third result below.

Result 3. (Test of Hypothesis 1). *The fraction of honest-looking predictions in Planned treatment is significantly higher than that in Impulsive.*

Evidence. Figure 3 plots the cumulative distribution of p_i^{max} for all three treatments. The fraction of the honest-looking predictions (*i.e.*, probability of $p_i^{max} \leq 50\%$), is 95% (72 of 76) in

Planned, and only 66% (40 of 61 subjects) in Impulsive. The difference between the two treatments is statistically significant ($p < .000$, two-sided proportion test).

The fraction of honest-looking predictions in Planned suggests an almost universal preference for appearing honest. However, in Impulsive, where the intention revelation was delayed, people responded to this very subtle change by making predictions that were much less honest-looking. This allowed them a greater opportunity to cheat.

Result 3 becomes perhaps even more evident as we take a closer look at the distribution of the highest probability p_i^{max} . In Figure 3, the distributions of p_i^{max} of the Planned and the Control treatments largely overlap with each other, and both are dominated by p_i^{max} of the Impulsive treatment at every percentile. For example, at roughly 60th percentile, p_i^{max} is 0.38 in both Planned and Control, but 0.49 in Impulsive. At 80th percentile, it is 0.40 in Planned and Control, and a staggering 0.86 in Impulsive.

Figure 3: Cumulative Distribution of Highest Probability p_i^{max}

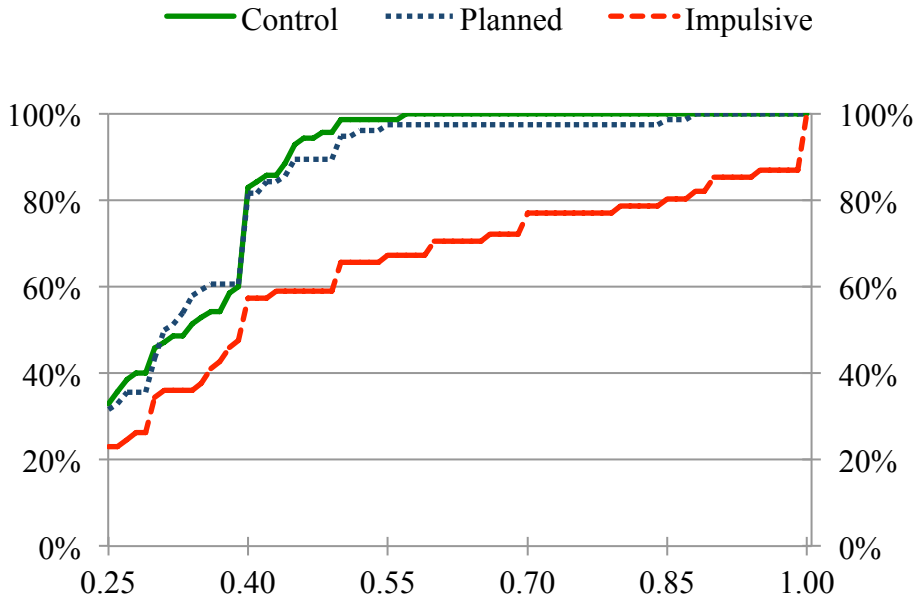


Table 1 summarizes some key information of p_i^{max} . For example, p_i^{max} starts at 25% (in objective predictions), and goes up to 57% in Control, 100% in Impulsive, and 88% in Planned.

It is worth noting that in Planned, no one felt comfortable pursuing the profit-maximizing strategy by predicting 100% on one single outcome, while several subjects did predict 100% in Impulsive. About 33% of subjects in Control made objective predictions, while only 23% of subjects in Impulsive did so. Interestingly, this fraction in Planned rose back up to 32%, almost identical to the Control treatment¹². The central tendencies also tell the same story, as the mean and median in Impulsive clearly stand above both Planned and Control treatments, while the latter two treatments are indistinguishable.

Table 1: Summary Statistics: Predictions

	Control	Impulsive	Planned
Number of observations	n=70	n=61	n=76
Fraction of objective predictions	33%	23%	32%
Max of p_i^{max}	57%	100%	88%
Mean (s.e.) of p_i^{max}	34% (1.0%)	51%(3.5%)	35% (1.4%)
Median of p_i^{max}	34%	40%	32%

Note: Standard errors are in parenthesis.

5.3. Preference for Being Honest

We examine the actuality of cheating behavior via prediction accuracy and earnings. Given that the dice are fair, we can draw inferences about cheating if subjects in one treatment make statistically more accurate predictions (and consequently higher earnings) than subjects in another treatment¹³. To measure prediction accuracy, we focus on highest-payoff outcomes and investigate whether they occur more often than what we expect from a fair die¹⁴. With a fair die, the expected frequency of a highest-payoff outcome is 25% if p_i^{max} is unique. When p_i^{max} is not

¹² Incidentally, these numbers are also the intercepts at highest probability=0.25 in Figure 1.

¹³ Planned is the only treatment in which subjects self-reported die roll outcomes, and we find that these self-reported outcomes differ significantly from the uniform distribution ($p=0.10$, chi-squared test). In contrast, die roll outcomes from Control or Impulsive do not differ from the objective distribution ($p=0.42$ and $p=0.60$, respectively, chi-squared test).

¹⁴ As objective predictions yield identical payoffs for all outcomes, they are excluded from this analysis.

unique, we adjust the expected frequency by the number of ties¹⁵. As summarized in Table 2, the expected frequency of highest-payoff outcomes is 32%, 30% and 29% in Control, Planned and Impulsive, respectively. They are very similar to each other, and we find no statistical difference in any pair-wise comparisons. Then, we compute the empirical frequency that subjects actually obtained highest-payoff outcomes, which is 36%, 71% and 85%, respectively. The difference between empirical and expected frequencies is significant in Planned and Impulsive, but not in Control, suggesting that cheating occurred in the first two treatments.

Table 2: Expected and Empirical Frequencies of Highest-payoff Outcomes
(Excluding Objective Predictions)

Highest-payoff outcomes	Control (n=47)	Planned (n=52)	Impulsive (n=47)
Expected Frequency	32%	30%	29%
Empirical Frequency	36%	71%	85%
Equality Test	p = 0.588	p < 0.001	p < 0.001

Notes: All p-values are obtained via the proportion test (two-sided).

The *expected frequency of a highest-payoff outcome* according to a fair die is 25% if a prediction has a unique highest probability p_i^{max} ; otherwise, the expected frequency must be adjusted for ties at p_i^{max} . For example, for the prediction [32%, 32%, 20%, 16%], the expected frequency of highest-payoff outcomes is 50%. The *empirical frequency of highest-payoff outcomes* is the fraction of subjects who actually reported that they obtained the highest-payoff outcome in the second round.

Subjects' earnings, due to their positive correlation with prediction accuracy, also reflect the differences between treatments. Recall that in Planned or Impulsive, a profit-maximizer could earn the maximum profit of \$25 by making sure that the prediction and outcome matched perfectly. In Control, however, subjects could only maximize expected earnings by submitting an objective prediction (25% for each outcome), for a guaranteed profit of \$15.63, regardless of the outcome. Table 3 summarizes what subjects actually earned in the experiment. The maximum earnings are \$25 in Impulsive, \$24.76 in Planned, and \$20.81 in Control. The medians are nearly

¹⁵ For example, if a subject predicts [32%, 32%, 20%, 16%], it is expected that the highest-payoff outcome (either 1 or 2) turns up with probability 50%. Hence, the expected frequency is 25% multiplied by the number of ties at p_i^{max} .

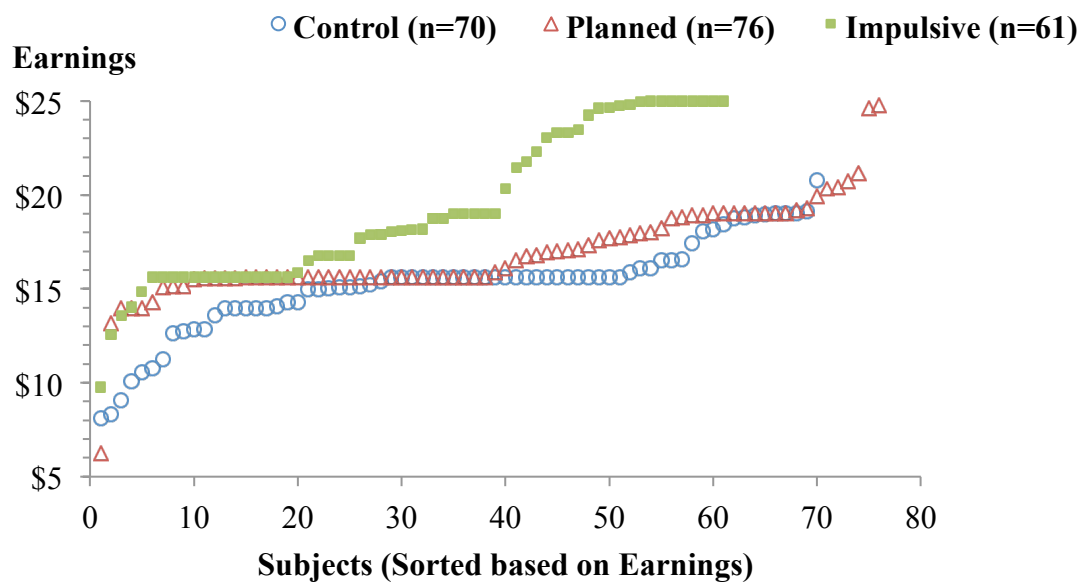
identical between Control and Planned, at \$15.63 and \$15.75, respectively, and clearly higher in Impulsive, at \$18.15. Figure 4 shows individual earnings, sorted from the lowest to the highest within each treatment.

Table 3: Earnings

	Control (n=70)	Planned (n=76)	Impulsive (n=61)
Min	\$8.12	\$6.25	\$9.75
Max	\$20.81	\$24.76	\$25.00
Median	\$15.63	\$15.75	\$18.15
Mean	\$15.26 (\$0.30)	\$16.90 (\$0.29)	\$19.15 (\$0.52)

Note: Standard errors are in parenthesis.

Figure 4: Scatter Plot of Earnings (Sorted from Low to High)



Comparing prediction accuracy and earnings between treatments, we present Result 4 and Result 5 below.

Result 4. (Test of Hypothesis 2). *Cheating occurred in both Planned and Impulsive where subjects made significantly (i) more accurate predictions and (ii) more money than those in Control.*

Evidence (i). The empirical frequency of highest-payoff outcomes subjects obtained was significantly higher than the expected frequency (adjusted by ties) in both Planned (71% vs. 30%, $p < 0.001$, two-sided proportion test) and Impulsive (85% vs. 29%, $p < 0.001$, two-sided proportion test) treatments. By contrast, predictions in Control were not more accurate than expected (36% vs. 32%, $p=0.588$, two-sided proportion test). These results suggest that significant cheating occurred in Planned and Impulsive where subjects had the opportunity to do so.

Evidence (ii). Average earnings in Planned were \$16.90, significantly higher than \$15.26 in Control ($p<0.00$, two-sided t -test). Average earnings in Impulsive were \$19.15, significantly higher than in Control ($p<0.00$, two-sided t -test) treatments. These results confirm Result 2, which finds that significant cheating occurred in Planned and Impulsive, where subjects had the opportunity to do so.

Result 5. (Test of Hypothesis 3). *More cheating occurred in Impulsive than in Planned, as subjects made significantly more accurate predictions in Impulsive than those in Planned.*

Evidence (i). Subjects' predictions in Impulsive were significantly more accurate than those in Planned (85% vs. 71%, $p < 0.095$, two-sided proportion test)¹⁶, suggesting that more cheating occurred in the Impulsive treatment. This result holds despite the fact that the expected frequency in the Impulsive treatment was slightly lower than that of Planned, *i.e.*, 29% vs. 30%, respectively.

Evidence (ii). Shown in Table 3, average earnings in Impulsive were significantly higher than those in Planned, at \$19.15 and \$16.90, respectively ($p<0.00$, two-sided t -test). Figure 4 shows that earnings in Impulsive are uniformly greater than those in Planned, and earnings in Planned again greatly exceed those in Control.

¹⁶ Generally tests on binary data are not powerful, and with our moderate sample size, a 10% significance level is already impressive.

5.4. Incomplete Cheating

In this subsection we investigate whether the incomplete cheating behavior reported in the literature occurs as a result of intrinsic preference for cheating partially, or as a consequence of preserving an honest appearance.

We first define a set of types, including honest, incomplete cheating, and maximum cheating. We then econometrically choose the optimal mixture from these pre-specified types that best characterizes subjects' decisions. If incomplete cheating is (not) included in this optimal mixture, we conclude that incomplete cheating is (not) an intrinsic preference for being honest.

Before we proceed, we verify the existence of truth-tellers in the population in Result 6.

Result 6: *A significant number of people reported truthfully.*

Evidence: First, almost one third of subjects in Planned submitted objective predictions, suggesting that many people follow the truth-telling strategy. Moreover, seven out of the 52 non-objective predictions reported outcomes corresponding to their lowest payoff in the second stage (13.5%)¹⁷. This result indicates that these people are intrinsically averse to cheating even when not being monitored.

Results 4 to 6 reveal that there exists a mixture of types in our population: some people are truth-tellers, while others are cheating in some way, either partially or maximally. We define these three types of intrinsic preferences for honesty as: (i) “truth-telling”; (ii) “maximum cheating”; and (iii) “one-step cheating.” The “truth-telling” type describes dogmatic truth tellers who report truthfully regardless of whether they are monitored. The “maximum cheating” type characterizes people who suffer little psychic disutility from cheating, and thus always report an outcome corresponding to the maximum profit. Finally, the “one-step cheating” type deviates from truth-telling, but only partially cheats by reporting the next available payoff level, which is not always the highest payoff¹⁸.

¹⁷ We adopt the common assumption in the literature that people would not cheat for worse outcomes.

¹⁸ The number of payoff levels in a prediction varies according to the number of ties. In particular, the “one-step cheating” type is identical to the “maximum cheating” type when there are only two payoff levels.

Type 1: (Truth-telling): The subject truthfully reports the die roll outcome, which follows the objective uniform distribution.

Type 2: (Maximum Cheating): The subject reports an outcome corresponding to the highest payoff.

Type 3: (One-step Cheating): The subject reports an outcome one payoff level higher than the actual realized outcome. In particular, the subject's strategy is always to report the highest-payoff outcome if he/she obtained such an outcome, and otherwise report an outcome corresponding to the next higher payoff level in relation to his/her realized outcome.

Note that we use the “one-step cheating” type to model the incomplete cheating behavior as an intrinsic preference not to deviate “too much” from honesty (see *e.g.*, Mazar et al. 2008; Lundquist et al. 2009). If people hold such preferences, then “one-step cheating” would explain self-reported die rolls better than the “maximum cheating” type.

Recall that the goal of this analysis is to determine which mixture of these three types best characterizes our subjects. As we have only one observation per subject, our inferences are based on aggregates that can be analyzed using a variant of the widely-used El-Gamal and Grether (1995) algorithm (see, *e.g.*, Anderson and Putterman 2006; Holt 1999; Houser and Winter 2004).

Allowing an error rate ε that is the same for all subjects, we say that each subject follows his/her decision rule (*i.e.*, type) with probability of $1 - \varepsilon$; with probability of ε , he/she trembles and reports all outcomes equally likely. Importantly, our “truth-telling” type also reports all outcomes equally likely due to the fact that the objective distribution is uniform. This implies two important features: (i) that the error rate ε is interpreted as the fraction of “truth-tellers” in the population; and (ii) that the “truth-telling” type is implicitly built into each mixture.

Before we specify the components of the likelihood function, we define the following notations. Let M_i denote the number of distinct payoff levels given by subject i 's prediction; rank all payoff levels from the lowest to highest. Let $t_i^{(j)}$ (where $j \in \{1, \dots, M_i\}$) be the number of ties at the j th lowest payoff level, so we have $\sum_{j=1}^{M_i} t_i^{(j)} = 4$ and that $t_i^{(M_i)}$ is the number of ties at the

highest payoff level. The indicator $D_i^{(j)}$ is 1 if subject i 's reported outcome corresponds to her j th lowest payoff level, and 0 otherwise; thus, we have $\sum_{j=1}^{M_i} D_i^{(j)} = 1$ and that $D_i^{(M_i)}$ indicates whether subject i 's reported outcome corresponds to the highest payoff.

Consider first the mixture of “truth-telling” and “maximum cheating” types. With probability $1 - \varepsilon$, a subject reports the highest-payoff outcome (“maximum cheating”); with probability ε , he/she reports each of the four outcomes with equal probability of 25% (“truth-telling”). This implies the following likelihood function (adjusted for ties) for the mixture of “maximum cheating” and “truth-telling” types.

$$\prod_{i=1}^n L_i(t_i^{(M_i)}, D_i^{(M_i)}) = \prod_{i=1}^n \left(1 - \varepsilon + \frac{\varepsilon}{4} t_i^{(M_i)}\right)^{D_i^{(M_i)}} \left(\frac{\varepsilon}{4} (4 - t_i^{(M_i)})\right)^{1 - D_i^{(M_i)}}$$

Next, consider the “one-step cheating” type, which predicts that subjects report outcomes corresponding to the next higher payoff level in relation to their realized outcomes. In particular, (i) the highest-payoff outcome is reported with the objective probability of obtaining the top two highest payoff levels; (ii) the lowest-payoff outcome is never reported; and (iii) the intermediate-payoff outcomes (which exist when $M_i > 2$) are reported with the objective probability of obtaining outcomes from the one-step lower payoff level. Adjusting for ties, we obtain the following likelihood function for the mixture of “one-step cheating” and “truth-telling”:

$$\prod_{i=1}^n L_i(t_i^{(j)}, D_i^{(j)}) = \begin{cases} \prod_{i=1}^n \left((1 - \varepsilon) \frac{t_i^{(M_i)} + t_i^{(M_i-1)}}{4} + \frac{\varepsilon}{4} t_i^{(M_i)}\right) & \text{if } D_i^{(M_i)} = 1 \\ \prod_{i=1}^n \left((1 - \varepsilon) \frac{t_i^{(j-1)}}{4} + \frac{\varepsilon}{4} t_i^{(j)}\right) & \text{if } D_i^{(j)} = 1 \text{ and } 2 \leq j \leq M_i - 1 \\ \prod_{i=1}^n \frac{\varepsilon}{4} t_i^{(1)} & \text{if } D_i^{(1)} = 1 \end{cases}$$

Finally, we consider the mixture of all three types, and obtain the likelihood as follows.

1. For each individual i , calculate the likelihoods for both mixtures: “maximum cheating and truth-telling” and “one-step cheating and truth-telling”; find the highest likelihood; and

2. Multiply the obtained highest likelihood across all n individuals, and find its maximum value by choosing the frequency of “truth-telling” ε .

To select among the three mixtures, we must include a penalty that increases with k , the number of types in the mixture. Following El-Gamal and Grether (1995, pp.1140-1141), our penalty is an uninformative “prior” distribution consisting of three parts. The first term is the prior for having k decision rules: $\frac{1}{2^k}$. The second term is the prior for selecting any k tuple of decision rules out of the universe of three decision rules: $\frac{1}{3^k}$. The third term says that each individual is assigned to one of the k decision rules independently and with equal probability: $1/k^n$.

Hence, our posterior mode estimates are obtained by maximizing the following:

$$\log\left(\prod_{i=1}^n \max L_i(t_i^{(j)}, D_i^{(j)})\right) - k \log(2) - k \log(3) - n * \log(k)$$

Table 4 reports the result of our analysis using data from Planned and Impulsive, and our final result follows.

Result 7: *The mixture of “maximum cheating” and “truth-telling” best characterizes subjects’ preference for being honest.*

Evidence. As reported in Table 4, the posterior mode is maximized with the mixture of “maximum cheating” and “truth-telling” in both Planned and Impulsive treatments. Take the Planned treatment for example: the readily calculated posterior odds ratio suggests that the mixture of “maximum cheating” and “truth-telling” types is about 7.5 times more likely than the mixture of “one-step shading” and “truth-telling.” Moreover, the estimates suggest that in the Planned treatment, 44% of subjects followed the “truth-telling” strategy. This rate of truth-telling is in line with what has been suggested in previous studies, including 39% reported by Fischbacher and Heusi (2008) and 51% reported by Houser et al. (2010). As expected, the fraction of truth-telling in Impulsive is lower (22%) than in Planned, and also somewhat lower than found in previous research. This may suggest that participants in previous studies were more concerned about the appearance of dishonesty than participants in our Impulsive treatment.

Table 4: Type Selection

	“Truth-telling” and “One-step Cheating”	“Truth-telling” and “Maximum Cheating”	“Truth-telling,” “One-step Cheating” and “Maximum cheating”
Planned (n=76)			
Number of Types	2	2	3
Posterior Mode	-128.98	-126.96	-143.18
Frequency of Truth-tellers	34%	44%	28%
Impulsive (n=61)			
Number of Types	2	2	3
Posterior Mode	-91.66	-88.12	-103.66
Frequency of Truth-tellers	14%	22%	12%

Willingness-to-Pay for an Honest Appearance. Finally, we compute subjects’ willingness-to-pay for appearing honest, using the money left on the table by the maximum cheaters in the Planned treatment. Subjects who reported highest-payoff outcomes, but did not predict certainty on a single outcome, nonetheless gave up a large amount of profit. Presumably, they did so to preserve an appearance of honesty. We measure this willingness-to-pay by the difference between earnings of subjects who reported the highest-payoff outcomes and the maximum profit. We find that average earnings by highest-payoff outcome reporters (n=36) are \$18.81, 75% of the maximum profit \$25. *Despite these subjects’ willingness to lie, they voluntarily left a quarter of their potential earnings on the table, suggesting a significant willingness-to-pay to appear honest.*

6. Discussion

The close resemblance in the probability predictions between Planned and Control demonstrates a prevalent preference for appearing honest. However, the appearance of honesty depends critically on perceptions, so any subtle shift in perceptions could lead to significant

changes in behavior. This is verified by our result, as cheating is significantly deterred when the intent to cheat must be revealed *before* (Planned) rather than *after* (Impulsive) a potentially dishonest act. We speculate that subjects in Impulsive might have strategically chosen to forget their original predictions, and persuaded themselves that they were indeed exceptional at predicting die rolls¹⁹. This type of rationalization and justification is consistent with Mazar et al.'s (2008) suggestion that people cheat only up to the level where they can benefit from cheating, but do not have to negatively update their self-image.

Even more striking is that independent evaluators also perceived the difference between Planned and Impulsive, despite its subtlety. This finding provides external validation to the difference in cheating behavior between Planned and Impulsive, and suggests that people do take into account outsiders' perspectives when they make the tradeoff between maintaining an honest appearance and profiting from cheating.

Finally, our type-classification analysis suggests that the “incomplete cheating” behavior observed in many previous studies can be attributed more to a preference for maintaining appearances than to an intrinsic aversion to maximum cheating. The majority of our subjects took the opportunity to signal an honest appearance. Having established that honest appearance, their second-stage decisions regarding the intrinsic preference for honesty can be better characterized as “maximum cheating” rather than “one-step cheating” (Result 5).

Two concerns regarding the subjects' understanding of the experiment deserve discussion. The first is that the Planned group could have been worried about the possibility and consequences of being caught cheating. To minimize this possibility, we emphasized to the subjects three times in the instructions that the die rolls would be completely private. We also encouraged them to roll many times, but to only report the first roll. This instruction made it more transparent that the chance of being caught was minimal. Moreover, subjects' answers in ex-post surveys indicated a clear awareness that they would be able to cheat²⁰. Finally, we also note that Ting's (2012) treatment comparison reports similar results to ours. Hence, we are confident that this issue did not significantly influence subjects' behavior.

²⁰ For example, when asked about their strategies for predictions, several subjects indicated that they were trying to maximize earnings.

The second concern is that subjects could be unaware of the earnings-maximizing strategy of assigning 100% to one outcome. Nonetheless, this seems unlikely to have been the case. The reason is that all participants were required to complete a quiz that assessed whether they understood how to assign probabilities to achieve maximum earnings. Participants could not continue the experiment without demonstrating this knowledge.

An alternative explanation for our results is that subjects' made their first-stage predictions not to signal an honest appearance, but to restrict the size of their lies in the second stage. This explanation seems improbable in light of the fact that probabilistic predictions in Planned are statistically identical to those in Control. It would be strikingly coincidental if decisions implied by maximum cheating-averse preferences were exactly consistent with preferences revealed under monitoring, as our data would require when combined with this alternative explanation.

Our results contribute to the literature by providing a unified explanation for a variety of behaviors reported in previous studies. For example, Mazar et al. (2009, p. 642) argued that subjects managed a positive self-view even after cheating. Their evidence was that there was no difference between the self-reported sense of honesty before and after a task in which subjects clearly cheated. However, our results suggest subjects, due to their desire to maintain an honest appearance, would have deliberately reported the same level of honesty, *especially after they cheated*.

Our results also have important implications for empirical research involving belief elicitation. Examples include field experimental studies and survey research (see, e.g., Manski 2004; Bellemare et al. 2008), where it is common to use non-saliently-rewarded procedures to elicit beliefs. A reason some choose not to use salient rewards is the “verification problem” (e.g., Manski 2004, footnote 11)²¹. The idea is that incentive-compatible mechanisms (such as the quadratic scoring rule) pay according to the outcome of the event, but realized outcomes are often difficult to verify in survey research. Hence, when the investigator relies on respondents'

²¹ Another reason people may not use scoring rules in large-scale surveys is that the level of cognitive ability required to understand the mechanisms is high. However, procedures for accurate belief elicitation have been developed and assessed within populations that include naïve respondents (Hao and Houser, 2011). On the other hand, recent evidence suggests belief elicitation accuracy is highly context specific, so that a generically “optimal” approach to belief elicitation may not exist (Armantier and Treich, 2010).

self reports, a sophisticated individual could maximize elicitation payoffs by first skewing her probabilistic prediction and then misreporting an outcome to match the prediction perfectly.

However, our results suggest that participants might not bias their predictions, even for those outcomes that cannot be verified. If this is true, then the elicited probabilities might in fact still hold value for out-of-sample inferences. This could weigh in favor of using incentivized approaches for belief elicitation, especially in light of the repeatedly demonstrated value of incentives for increasing participant attention and focus on the task (see, *e.g.*, Smith 1965; Houser and Xiao 2011).

In conclusion, our results lend support to the view that incentivized mechanisms can be used to elicit beliefs even when event outcomes cannot be verified. The reason is a ubiquitous preference for an honest appearance. The consequence is that when the outcome is not verifiable, in-sample predictions may be too “accurate”: 71% subjects in Planned reported that they indeed “obtained” the outcome they predicted was most likely to occur, a statistically significant departure from the expected frequency of 30% (Table 2). On the other hand, the predictions do well out-of-sample in the sense that the first-stage predictions in Planned are identical to those in Control.

One limitation of this study is that its results might depend on the specific payoffs we employed. Whether people are less concerned about their honest appearance when monetary incentives are sufficiently high is one important testable hypothesis left for future studies. Also important is to understand how willingness to cheat varies across people and settings. In particular, understanding how demographics, religious and political views affect the rationalization of cheating behavior might shed light on institutions that can mitigate terrorism, misconduct and corruption.

References:

- Akerlof, George A. 1983. "Loyalty Filters." *American Economic Review*, 73(1): 54-63
- Andersen, Steffen, John Fountain, Glenn . Harrison, and Rutström, E. Elisabet. 2010. "Estimating Subjective Probabilities." Working Paper 2010-06, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.
- Anderson, Christopher and Louis Putterman. 2006. "Do Non-Strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism." *Games and Economic Behavior*, 54(1): 1-24
- Andreoni, James and Bernheim, Douglas. 2009. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*, 77(5): 1607-1636
- Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review*, 99(1): 544-555
- Armantier, Olivier and Nicolas Treich. 2010. "Eliciting Beliefs: Proper Scoring Rules, Incentives, Stakes and Hedging." Working paper.
- Bellemare, Charles, Sabine Kroger, and Arthur Van Soest. 2008. "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities." *Econometrica*, 76(4): 815-839
- Bénabou, Roland, and Jean Tirole. 2004. "Willpower and Personal Rules." *Journal of Political Economy*. 112 (4): 848-886
- Bodner, Ronit, and Drazen Prelec. 2003. "Self-Signaling and Diagnostic Utility in Everyday Decision Making." *The Psychology of Economic Decisions*, vol. 1, edited by Isabelle Brocas and Juan D. Carrillo. Oxford: Oxford Univ. Press.
- Buccioli, A., and M. Piovesan. 2009. "Luck or Cheating? A Field Experiment on Honesty with Children." mimeo.
- Cornish, D. B., & Clarke, R. V. G. (Eds.). (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York, NY: Springer-Verlag.
- El-Gamal, Mahmoud and Grether, David M. 1995. "Are People Bayesian? Uncovering Behavioral strategies." *Journal of the American statistical Association*. 90 (432): 1137-1145
- Ellingsen, Tore and Johannesson, Magnus. 2004. "Promises, Threats and Fairness." *Economic Journal*, 114 (495): 397-420
- Felson, M. (2002). *Crime and everyday life* (3rd ed.). Thousand Oaks, CA: Pine Forge.
- Felson, Richard B. and Michael Massoglia. 2011. "When is Violence Planned?" Forthcoming, *Journal of Interpersonal Violence*.
- Fischbacher, U., and F. Heusi. 2008. "Lies in Disguise – An experimental study on cheating." Thurgau Institute of Economics Working Paper.
- Gneezy, Uri. 2005. "Deception: The Role of Consequences." *American Economic Review*, 95(1): 384-394

- Greene, Joshua D. and Joseph M. Paxton. 2009. "Patterns of Neural Activity Associated with Honest and Dishonest Moral Decisions." *Proceedings of National Academy of Science*, 106(30): 12506-12511
- Hannan, R.Lynn, Frederick . Rankin and Kristy L. Towry. 2006. "The Effect of Information System on Honesty in Managerial Reporting: A Behavioral Perspective." *Contemporary Account Research* 23(4): 885-918
- Hao, Li and Daniel Houser. 2011. "Belief Elicitation in the Presence of Naïve Respondents: An Experimental Study." Working Paper.
- Holt, Debra. 1999. "An Empirical Model of Strategic Choice with an Application to Coordination Games." *Games and Economic Behavior*, 27: 86-105
- Houser, Daniel and Erte Xiao. 2011. "Classification of Natural Language Messages Using a Coordination Game." *Experimental Economics*, 14(1): 1-14
- Houser, Daniel, Stefan Vetter, and Joachim Winter. 2010. "Fairness and Cheating." Working paper.
- Houser, Daniel and Joachim Winter. 2004. "How Do Behavioral Assumptions Affect Structural Inference? Evidence from a Laboratory Experiment." *Journal of Business and Economic Statistics*, 22(1): 64-79.
- Hurkens, S. and N. Kartik. (forthcoming). "Would I Lie to You? on Social Preferences and Lying Aversion." *Experimental Economics*.
- Jiang, Ting. (2012). "The Mind Game: Invisible Cheating and Inferable Intentions." Working paper.
- Lundquist, Tobias, Tore Ellingsen, Erik Gribbe and Magnus Johannesson. 2009. "The Aversion to Maximum cheating." *Journal of Economic Behavior & Organization*, 70: 81-92
- Manski. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329-1376
- Mazar, Nina; Amir, On and Ariely, Dan. 2008. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." *Journal of Marketing Research*, 45: 633-644
- Nyarko, Yaw, and Andrew Schotter. 2002. "An Experimental Study of Belief Learning Using Elicited Beliefs." *Econometrica*, 70 (3): 971-1005
- Rode, J. 2010. "Truth and Trust in Communication: an Experimental Study of Behavior under Asymmetric Information", *Games and Economic Behavior*, 68 (1): 325–338.
- Sánchez-Pagés, S. and Vorsatz, M. 2007. "An experimental study of truth-telling in a sender-receiver game." *Games and Economic Behavior*, 61: 86–112.
- Sánchez-Pagés,, S. and Vorsatz, M. 2009. "Enjoy the silence: an experiment on truth-telling." *Experimental Economics*, 12: 220–241
- Savage Leonard J. 1971. "Elicitation of Personal Probabilities and Expectations." *Journal of American Statistical Association*, 66(336): 783-801
- Spence, A. M. 1974. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge: Harvard University Press.

- Sutter, M. 2009. "Deception through Telling the Truth?! Experimental Evidence from individuals and Teams." *Economic Journal*, 119: 171-175
- Smith, Vernon. 1965. "Effect of Market Organization on Competitive Equilibrium." *Quarterly Journal of Economics*, 78(2): 181-201
- Williamson, O. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Winkler, Robert L. 1969. "Scoring Rules and the Evaluation of Probability Assessors." *Journal of American Statistical Association*, 64(327): 1073-1078

Appendix 1: Instructions and Decision Sheets.

(Control Treatment)

Instructions

Welcome to this experiment. In addition to \$5 for showing up on time, you will be paid in cash based on your decisions in this experiment. Please read the instructions carefully. No communication with other participants is allowed during this experiment. Please raise your hand if you have any questions, and the experimenter will assist you.

There are two parts to this experiment. In Part I, you predict the percent chance that you will roll a one, two, three or four on a single roll of a fair four-sided die. In Part II, you will roll the die as many times as you wish on your own, and the experimenter will come to your desk and ask you to roll it exactly once. The experimenter will watch your roll and record the outcome.

Your earnings depend on how close your prediction in part I is to the outcome of your single roll in front of the experimenter in Part II. We'll now explain each part in more detail.

How You Earn Money:

In Part I, your task is to predict the percent chance that your single roll of a four sided die in front of the experimenter will be a one, two, three or four. *Your predictions are not what you hope will happen, but what you believe will happen.* Remember the die is “fair”, which means each number should be equally likely. However, sometimes people have a “hunch” that rolling one number is more likely, and if you have a hunch you should indicate this when you write down your percent chances.

You earn more money when your predictions are closer to the outcome of your single roll. For example, you earn the most if you predict 100% chance that you will roll a certain number, and then you actually do roll that number. On the other hand, you earn the least if you predict 100% chance you will roll some number, and then you don't roll that number.

Please use the spreadsheet tool on your computer to explore how different predictions affect your earnings depending on the number you roll in front of the experimenter. Below we provide a few examples.

If you predict percent chances:	If your single roll in front of the experimenter is	Your earnings are
1: 0%; 2: 100%; 3: 0%; 4: 0%	2	\$25.00
	4	\$0.00
1: 0%; 2: 50%; 3: 50%; 4: 0%	2	\$18.75
	4	\$6.25
1: 25%; 2: 25%; 3: 25%; 4: 25%	2	\$15.63
	4	\$15.63
1: 20%; 2: 20%; 3: 20%; 4: 40%	2	\$14.00
	4	\$19.00
1: 33%; 2: 0%; 3: 34%; 4: 33%	2	\$8.33

	4	\$16.58
--	---	---------

To summarize, the formula for calculating your earnings is

*Earnings = \$12.5 + \$25 * percent chance you predicted for the outcome of your single roll in front of the experimenter*

*–\$12.5 * (percent chance you predicted for number 1)²*

*–\$12.5 * (percent chance you predicted for number 2)²*

*–\$12.5 * (percent chance you predicted for number 3)²*

*–\$12.5 * (percent chance you predicted for number 4)²*

This formula shows how exactly your earnings are calculated. To understand how different predictions affect your earnings, please take your time and use the spreadsheet on your computer terminal to explore.

After you submit your decision sheet, the experimenter will come to your table and ask you to roll the die exactly once, and record the number. This number will determine your earnings.

Decision Sheet:

Your experiment ID: _____

Print your name: _____

Part I:

Before you roll the four-sided die, please predict the percent chance that each number turns up of your single roll in front of the experimenter.

A. What is the percent chance that the number 1 turns up? _____

B. What is the percent chance that the number 2 turns up? _____

C. What is the percent chance that the number 3 turns up? _____

D. What is the percent chance that the number 4 turns up? _____

Important: the sum of all four answers MUST be 100%. Is the total of your four answers equal to 100%? Please circle one of the following: Yes No

(Planned Treatment)

Instructions

Welcome to this experiment. In addition to \$5 for showing up on time, you will be paid in cash based on your decisions in this experiment. Please read the instructions carefully. No communication with other participants is allowed during this experiment. Please raise your hand if you have any questions, and the experimenter will assist you.

There are two parts to this experiment. In Part I, you predict the percent chance that you will roll a one, two, three or four on your first roll of a fair four-sided die. **In Part II, you will roll the die as many times as you wish on your own, but you will need to remember the first number you roll.**

Your earnings depend on how close your prediction in part I is to the outcome of your first roll in Part II. We'll now explain each part in more detail.

How You Earn Money:

In Part I, your task is to predict the percent chance that your first roll of a four sided die will be a one, two three or a four. *Your predictions are not what you hope will happen, but what you believe will happen.* Remember the die is “fair”, which means each number should be equally likely. However, sometimes people have a “hunch” that rolling one number is more likely, and if you have a hunch you should indicate this when you write down your percent chances.

You earn more money when your predictions are closer to the outcome of your first roll. For example, you earn the most if you predict 100% chance that you will roll a certain number, and then you actually do roll that number. On the other hand, you earn the least if you predict 100% chance you will roll some number, and then you don't roll that number.

Please use the spreadsheet tool on your computer to explore how different predictions affect your earnings depending on the first number you roll. Below we provide a few examples.

If you predict percent chances:	If your first roll is	Your earnings are
1: 0%; 2: 100%; 3: 0%; 4: 0%	2	\$25.00
	4	\$0.00
1: 0%; 2: 50%; 3: 50%; 4: 0%	2	\$18.75
	4	\$6.25
1: 25%; 2: 25%; 3: 25%; 4: 25%	2	\$15.63
	4	\$15.63
1: 20%; 2: 20%; 3: 20%; 4: 40%	2	\$14.00
	4	\$19.00
1: 33%; 2: 0%; 3: 34%; 4: 33%	2	\$8.33
	4	\$16.58

To summarize, the formula for calculating your earnings is

$$\begin{aligned} \text{Earnings} = & \$12.5 + \$25 \\ & * \text{percent chance you predicted for the outcome of your first roll} \\ & -\$12.5 * (\text{percent chance you predicted for number 1})^2 \\ & -\$12.5 * (\text{percent chance you predicted for number 2})^2 \\ & -\$12.5 * (\text{percent chance you predicted for number 3})^2 \\ & -\$12.5 * (\text{percent chance you predicted for number 4})^2 \end{aligned}$$

This formula shows how exactly your earnings are calculated. To understand how different predictions affect your earnings, please take your time and use the spreadsheet on your computer terminal to explore.

After you submit your Decision Sheet 1, please take your time and roll the die as many times as you wish on your own. You will need to remember and report the first number you rolled, because this number will determine your final earnings.

Decision Sheet 1:

Your experiment ID: _____

Print your name: _____

Part I:

Before you roll the four-sided die, please predict the percent chance that each number turns up for your first roll.

A. What is the percent chance that the number 1 turns up? _____

B. What is the percent chance that the number 2 turns up? _____

C. What is the percent chance that the number 3 turns up? _____

D. What is the percent chance that the number 4 turns up? _____

Important: the sum of all four answers **MUST** be 100%. Is the total of your four answers equal to 100%? Please circle one of the following: Yes No

Decision Sheet 2:

Your experiment ID: _____

Part II:

Now take your time and roll the die as many times as you wish on your own, but please remember the first number you roll, because you will need to write it down and it will determine your earnings:

Please write down the first number you rolled: _____

(Evaluation Group)

Instructions

Welcome to this experiment. In addition to \$5 for showing up on time, you will be paid in cash based on your decisions in this experiment. Please read the instructions carefully. No communication with other participants is allowed during this experiment. Please raise your hand if you have any questions, and the experimenter will assist you.

Summary:

Instructions of two previous experiments have already been handed to you. We denote these as Experiment 1 and Experiment 2. *Your task today is to evaluate whether a participant in these previous experiments **appeared to have cheated**.*

Before reading further here, please turn to and read the instructions from the previous experiments.

CONTINUE HERE AFTER READING THE INSTRUCTIONS FOR EXPERIMENT 1 AND 2.

Remember, you are encouraged to review those instructions as you make your decisions today. Understanding the environments of the previous experiments will help you earn more money during your task today.

In the previous two experiments, the maximum amount a participant can earn is \$25, in addition to \$5 for showing up on time. Here is how they can earn that maximum amount.

In Experiment 1, participants can earn the maximum \$25 if they predict outcome 1 (, 2, 3, or 4) will occur with 100%, and then report that their private roll was indeed 1 (, 2, 3, or 4, respectively). Of course, nobody *except* themselves knows which outcome they actually rolled.

In Experiment 2, participants do not reveal their prediction until after they observe the public die roll outcome. They are asked to report their original prediction, and they can earn the maximum \$25 by reporting that they predicted that it was 100% certain that the outcome would be what they publicly rolled. Of course, nobody *except* themselves knows their original prediction.

In general, the more probability the participants predict on any single outcome, the closer are their earnings to the maximum profit \$25. Because the highest probability prediction indicates the highest earnings, in the following questions, we will focus on the highest probability that the participants predicted on the die roll outcomes.

Your task:

Your task today is to evaluate whether a participant in these previous experiments appeared to have cheated. There are a total of 18 questions, divided evenly into three sets, A, B and C. Questions in Set A describe decisions from six scenarios that occurred in Experiment 1. Questions in Set B consist of six scenarios from Experiment 2. The final six questions in Set C ask you to compare between Experiments 1 and 2, and evaluate in which scenario the participant appeared more likely to have cheated.

Your earnings

We will randomly select three questions, one from each set of questions. Your answer to each of those questions will be compared with the answers of the other people in this room. If your answer agrees with the most common answer of all participants in this room, you will earn \$5 for that question. Therefore, you can earn up to \$15 in addition to the \$5 for showing up today, depending on how well your answers match the answers of the other people in this room.

Note: You will be given one set of questions at a time. Once you submit your answers, you won't be able to change them.

Decision Set A

*Evaluate the following scenarios that occurred in **Experiment 1**:*

Scenario 1: In Experiment 1, a participant first revealed his or her prediction that outcome **2** would occur with **50%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **2**.

Question 1: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Scenario 2: In Experiment 1, a participant first revealed his or her prediction that outcome **3** would occur with **60%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **3**.

Question 2: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Scenario 3: In Experiment 1, a participant first revealed his or her prediction that outcome **1** would occur with **40%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **1**.

Question 3: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Scenario 4: In Experiment 1, a participant first revealed his or her prediction that outcome **4** would occur with **25%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **4**.

Question 4: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Scenario 5: In Experiment 1, a participant first revealed his or her prediction that outcome **3** would occur with **85%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **3**.

Question 5: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Scenario 6: In Experiment 1, a participant first revealed his or her prediction that outcome **2** would occur with **100%** chance, and that was the highest probability the participant placed on any outcome. The participant then rolled the die *privately*, and reported that the die roll was **2**.

Question 6: Do you think the majority of people in this room believe that this participant cheated?

☐Yes ☐No

Decision Set B:

*Evaluate the following scenarios that occurred in **Experiment 2**:*

Scenario 7: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **2**. The participant then reported to have *privately* predicted that **2** would occur with **50%** chance, and that this was the highest probability the participant placed on any outcome.

Question 7: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Scenario 8: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **3**. The participant then reported to have *privately* predicted that **3** would occur with **60%** chance, and that this was the highest probability the participant placed on any outcome.

Question 8: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Scenario 9: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **1**. The participant then reported to have *privately* predicted that **1** would occur with **40%** chance, and that this was the highest probability the participant placed on any outcome.

Question 9: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Scenario 10: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **4**. The participant then reported to have *privately* predicted that **4** would occur with **25%** chance, and that this was the highest probability the participant placed on any outcome.

Question 10: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Scenario 11: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **3**. The participant then reported to have *privately* predicted that **3** would occur with **85%** chance, and that this was the highest probability the participant placed on any outcome.

Question 11: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Scenario 12: In Experiment 2, a participant first predicted *privately* the percent chance of each outcome, then rolled the die publicly and observed the outcome was **2**. The participant then reported to have *privately* predicted that **2** would occur with **100%** chance, and that this was the highest probability the participant placed on any outcome.

Question 12: Do you think the majority of people in this room believe that this participant cheated?
☐Yes ☐No

Decision Set C:

*The following questions compare scenarios that occurred in **Experiment 1** and in **Experiment 2**. Take Question 13 for example: if your answers in Scenarios 1 and 7 are different, say you answered “Yes” in Scenario 1 and “No” in Scenario 7, then effectively you already indicated that your answer for Question 13 is “Scenario 1.” Hence, in this set, you only need to answer those questions where you had the same answers before.*

Question 13: If your answers in **Scenarios 1 and 7** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 1 or in Scenario 7?

☐ Scenario 1 ☐ Scenario 7 ☐ Equally likely

Question 14: If your answers in **Scenarios 2 and 8** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 2 or in Scenario 8?

☐ Scenario 2 ☐ Scenario 8 ☐ Equally likely

Question 15: If your answers in **Scenarios 3 and 9** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 3 or in Scenario 9?

☐ Scenario 3 ☐ Scenario 9 ☐ Equally likely

Question 16: If your answers in **Scenarios 4 and 10** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 4 or in Scenario 10?

☐ Scenario 4 ☐ Scenario 10 ☐ Equally likely

Question 17: If your answers in **Scenarios 5 and 11** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 5 or in Scenario 11?

☐ Scenario 5 ☐ Scenario 11 ☐ Equally likely

Question 18: If your answers in **Scenarios 6 and 12** are the same, do you think the majority of people in this room believe that the participant more likely to have cheated in Scenario 6 or in Scenario 12?

☐ Scenario 6 ☐ Scenario 12 ☐ Equally likely

Appendix 2: Screenshot of Excel Tool for the Quadratic Scoring Rule

