Title
Essays in Behavioral and Experimental Economics

Author
Jiang, Xin

Publication Date
2024

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Essays in Behavioral and Experimental Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Xin Jiang

Committee in charge:

Professor Gary Charness, Chair, *in memoriam*
Professor Erik Eyster
Professor Ignacio Esponda
Professor Antony Millner

June 2024

The Dissertation of Xin Jiang is approved.

Professor Erik Eyster

Professor Ignacio Esponda

Professor Antony Millner

Professor Gary Charness, Committee Chair, *in memoriam*

June 2024

Essays in Behavioral and Experimental Economics

To the people in my sweet hometown Laiyang, whose perseverance inspires me. Your hard work shapes my understanding, your spirit fuels my aspirations. This dissertation is dedicated to you with heartfelt gratitude and immense love.

Acknowledgements

I would like to express my deepest gratitude to my Gary Charness, Erik Eyster, Ignacio Esponda for their unwavering professional guidance and support throughout my PhD journey. Their insights and expertise have been invaluable in shaping my research and academic growth. I am also deeply thankful to the Ryan Oprea, Sevgi Yuksel, Daniel Martin, Cheong-zhong Qin, John Hartman and Mark Patterson who have contributed significantly to my development, both in research and teaching. Additionally, I also want to thank my coauthors Gary Charness, Jing Zhou, Bohan Ye, and gradaute students at the research group who have spent countless hours discussing specific research steps and finding solutions to problem.

To my partner Chong Liu, I owe a special thanks for his endless support and belief in me. His encouragement has been a cornerstone of my perseverance, especially during the challenging times when I felt stuck or uncertain. His faith in my abilities has given me the strength to push forward and achieve my goals. I am equally grateful to my parents for their constant support and understanding. Their unwavering belief in my potential and your readiness to stand by me through every hard decision have made all the difference.

A special note of thanks goes to my committee chair Gary Charness, *in memoriam*, who has left an indelible mark on my life and work. His profound influence extended far beyond the realm of research. He taught me three invaluable lessons that I will carry with me always: keeping moving, making friends, and being famous. His wisdom on the importance of perseverance in the face of obstacles, the value of building meaningful relationships, and the pursuit of excellence and recognition in one's field have profoundly shaped both my academic and personal life. His legacy continues to inspire me every day, and I am forever grateful for his guidance and friendship.

Curriculum Vitæ

Xin Jiang

Education

2024	Ph.D. in Economics, University of California, Santa Barbara, California, USA.
2019	M.Phil. in Economics, The Chinese University of Hong Kong, Hong Kong.
2014	B.A. in Economics (Minor), Peking University, Beijing, China.
2014	B.E. in Mechatronic Engineering, Beijing Institute of Technology, Beijing, China.

Publications

- **Fortune and Identity.**
  Gary Charness and Xin Jiang.
  Economics Letters, 222(110954):1-3, 2023.

Work in Progress

- **Dynamic Binary Method in Belief Elicitation.**
  Xin Jiang and Jing Zhou.
- **Response Time in Deceptive Communication**
  Xin Jiang and Bohan Ye.

Awards

- **Outstanding Teaching Assistant Award**, UCSB, Spring 2024.
- **Humanities & Social Sciences Research Grant**, UCSB, Spring 2022.
- **Third Place in Econ Grad Slam Competition**, UCSB, Spring 2021.
- **Research Fellowship**, UCSB, Fall 2019.

Abstract

Essays in Behavioral and Experimental Economics

by

Xin Jiang

The dissertation consists of three experimental studies on economics, including in-group bias, deception heuristics, and belief elicitation.

Chapter 1 studies the root of in-group bias. Literature has found that social identities and minimal group paradigm can generate in-group bias, but seldom studies what non-label activities could generate group affiliation. I study the effect of common experience on group affiliation through a lab experiment. The results show that common fortune experience works, while common misfortune does not. These results violate results from previous studies, and suggest that some other perspectives work beyond pure in-group favoritism, for example, the sense of deservingness.

Chapter 2 studies the response time in lying detection. The inclusion of response time indicators has become a common feature in the contemporary landscape of social media sites. What private information does the response time carry when there is a conflict of interest, and do people use it to improve their welfare? We portray a model and design a modified cheap talk game to study the intricate interplay between response time, private information, and its influence on users' well-being, tailored to situations where truth discovery is time consuming. Our investigation uncovers a noteworthy sender hope to not have to lie to get what she wants. Given this preference, the private information reveals the consideration process, instead of the mechanical discovery process. We find that when there is an apparent conflict of interest, the longer the response time, the less credible the message. However, receivers are unable to extract substantial welfare gains

through the response time. Furthermore, when senders are aware of the availability of their response time, they are able to manipulate it.

Chapter 3 studies the belief elicitation method. Beliefs or perceptions play a central role in studying economic behavior, yet eliciting them accurately presents challenges. We introduce a novel elicitation method, called the *Dynamic Binary Method* (DBM), designed to address the common challenge individuals face in pinpointing the best point estimate of their beliefs, particularly when their beliefs are imprecise. Unlike Classical Methods (CM), which require respondents to make absolute judgments and form a point estimate of their true beliefs, DBM guides them through a series of binary relative judgments, enabling them to express interval beliefs by exiting the process at any step. To assess the empirical validity of DBM, we conduct both within-subject and between-subject experiments using a diverse range of perception tasks drawn from previous literature and CM as a benchmark of performances in each task. We find that DBM does not perform significantly differently from CM at the aggregate level, regardless of whether the perception questions use artificial/laboratory settings or real-life settings, and irrespective of the measurement used. Notably, DBM outperforms CM when the objective truth is extreme. Furthermore, we find a negative correlation between the length of stated beliefs in tasks using DBM and their accuracy. Additionally, we find that the length stated in DBM can predict respondents' performance in CM tasks at the aggregate level, albeit not strictly in a monotonic manner. Finally, we explore methods to use DBM-collected data for predicting stated point beliefs in DBM, offering insights into potential applications of the method beyond its immediate implementation.

JEL C92, D91, C70

Contents

Chapter 1

Fortune and identity

Joint work with Gary Charness

1.1 Introduction

Both psychologists and economists have shown evidence that people exhibit in-group bias in a variety of contexts, including tax preference, cooperation, punishment and truth telling. A number of studies have found in-group favoritism even in the trivial, ad hoc intergroup categorization, i.e. minimal group paradigm. However, when applying this concept in the real world, economists only focus on ingrained categories, like race, gender, ethnic and religion. I think it can be more popular than what have been found. For example, it is intuitive to think that common experience would establish a bond among people, and encourage people to help and cooperate with each other. Therefore, matching people with the same experience together might benefit people via in-group bias. However, it is hard to study the effect of common experience in the real world without confounds. Lab experiments are a better way to study the effect of common experiences in a highly-controlled (albeit stylized) environment.

This study aims to investigate how common experiences of misfortune and fortune shape individuals' sense of group affiliation. I focus on distribution decisions in this study, which is the most popular way to study in-group bias in identity literature, and leave other forms of interactions, like cooperation and strategic communication for future research. I focus on common experiences of fortune and misfortune because it is penetrated into all corners of life, from whether you are born rich or poor, whether you are born as a discriminated race, to whether you get into a university or a job that you're qualified for. I assume that both common experiences generate group cohesion. In addition, negative feelings from common misfortune are stronger than positive feelings from common fortune, I further assume that the unfortunate participants show more in-group favoritism than fortunate participants.

My experiment consists of two stages: a common experience generation stage and a distribution stage. In the common experience generation stage, subjects were assigned to two payment schemes. For the control allocators, subjects were paid a fixed fee for completion. For all other subjects, they were paid randomly due to a lottery result. This payment variation allows me to compare the effects of sharing common fortune and misfortune experience on distribution decisions in the next stage.

In the distribution stage, I asked subjects to make distribution decisions of extra money for two other recipients. Except for control recipients, each subject made decisions under three scenarios: if both recipients were fortunate, if both recipients were unfortunate, if one of the recipients was fortunate, and the other was unfortunate. My main question is whether an allocator's distribution decisions to others are affected by the common experience with recipients, and if so, how does it work.

The results do not support my hypotheses. The asymmetry of effects of common experiences works in the opposite way. Unfortunate allocators distribute similarly as the neutral allocators, the fortunate allocators distribute more to the fortunate allocators.

In this way, it seems that the fortunate allocators were affected by common experience, while the unfortunate allocators did not, which violates my hypothesis.

My experiment contributes to the literature studying the effect of pure luck on distribution decisions. My results are different from results found in the literature. When people know that their results are determined by luck, Espinosa et al. [2020] found that there was no big difference in redistribution decisions among different types of allocators, Cassar and Klein [2019] found lottery failures were more likely to favor other lottery failures than other people, and there was no big effect of common experience for lottery winners. The difference of decisions may come from the inequality generation stage. In both Cassar and Klein [2019], Espinosa et al. [2020], they asked subjects to do some real effort tasks, and informed them of their absolute or relative performance. In addition, subjects were informed that their final outcomes are actually randomly decided. It was not clear how people would identify themselves with two pieces of information. For example, if a person got a low score but a high payoff, would he build a bond with a low-payoff recipient for the reason that he thought he deserved low payoff, or with a high-payoff recipient simply for the same payoff? It is hard to identify the in-group bias with the ambiguous relationship between outcomes and identities. The largest contribution of my experiment is that I resolve this concern. Participants do the same task, and have the same performance. They spend several minutes experiencing the random generation process, so it is pretty clear that their payoffs are due to random luck only.

One explanation for for the asymmetric effects of common experiences is the asymmetric sense of deservingness. With the same performance, people who got above average would like to legitimize their payoffs, to persuade themselves and others that they deserve their payoffs. One way to legitimize is not to punish others who got the same payoff as them, or reward them in the extreme case. People who got below average have no such demand. So we see that "in-group bias" for fortunate people, but not for unfortunate people. This explanation can further illustrate the difference between my results and results in the literature. Cassar and Klein [2019], Espinosa et al. [2020] both informed participants of their performance in the task, which might be used as the way to support deservingness. So fortunate participants in their experiment did not need to provide additional evidence for legitimization, therefore, they might behave differently.

Another explanation is that all people have in-group bias, but the environment for unfortunate people is not optimal to exhibit such preference. We can separate these two mechanisms with future experiments. This explanation is consistent with the claim of Chen and Li [2009] that the sense of group affiliation from any arbitrary separation is common and symmetric.

The remainder of the paper is organized as follows. Literature review is discussed in Section 2. Section 3 describes the experimental design. Section 4 illustrates the conceptualization. Section 5 provides results. Section 6 concludes with discussion.

1.2 Literature Review

This paper relates to two major strands of literature. First, it contributes to the fast-growing experimental literature on distributional preferences. Many studies have found that people have fairness preferences not only for outcomes, but also for how those outcomes are reached (the source of inequality). With dictator games, Konow [2000] found that people are more inequality-averse when merit rather than luck is the source of inequality, and call this “accountability principle”. He provided evidence that people are more likely to reward individuals based on their efforts, and to compensate them for back luck. Cappelen et al. [2007, 2013] and Tinghög et al. [2017] found most people favor ex-post egalitarian fairness with respect to lucky and unlucky risk takers, but not with respect to risk-takers and participants who chose a safe alternative to avoid risk. Many studies have provided evidence that these findings are also observed in dictator games when self-interest concerns have been eliminated.

These studies are typically based on the hidden assumption that individuals have a fairness view that is fixed among the distributional situations in which they find themselves. In contrast, a sub-strand of literature on distributional preferences investigates whether people's fairness views are malleable in terms of an individual's economic experiences. When decision makers take their self-interest into consideration, there are consistent results that decision makers tried to maximize their ex-post payments regardless of economic status, and in terms of redistribution decisions, high-reward subjects redistribute substantially less than low-reward subjects [Almås et al., 2017, Durante et al., 2014, Gee et al., 2017, Lefgren et al., 2016].

Some studies tried to further explore if the magnitude of redistribution responds to income generation process, (i.e. if people redistribute more if the inequality is generated by luck than if generated by merit), but there are conflicting results. Durante et al. [2014] found that there was no systematic difference in redistribution rates between merit and luck income-determination methods. Gee et al. [2017] found that this was only true for high-reward participants. They found that low-reward and middle-reward participants redistributed more when the payoff was based on performance, which violates findings from much of the previous work that people supported for more redistribution when income is allocated randomly. Lefgren et al. [2016] found that both high-reward and low-reward participants redistributed more when the payoff was based on luck, which is consistent with much of the previous work. It may be true that self-interest makes it difficult to study the asymmetric effect of the fairness perspectives of different outcomes.

To clearly identify the experience effect on distributional preferences net of self interest, some papers consider dictator games where the dictator's pay is the same regardless of the choice made. With the manipulation of ambiguity as to whether success or failure should be attributed to merit or luck factors, Deffains et al. [2016] provided evidence that the experience of success or failure in a preceding real-effort task caused people's distributional preferences to diverge due to self-serving bias, i.e. successful participants are more likely to attribute their success to their effort rather than luck, and they opted for less redistribution according to Konow [2000]'s "accountability principle". Espinosa et al. [2020] provided further evidence that the belief of deservingness was the key for self-serving bias, they successfully debiased preferences in Deffains et al. [2016] by stressing the importance of chance in full separation of outcomes before the task.

In this literature, close to my work is Cassar and Klein [2019], who conducted a disinterested-dictator design where there was no ambiguity between merit- and luck-determination protocols. However, there are three main differences between their work and my experiment. First, in their work, there was some ambiguity in how participants defined themselves in their experiment. In their identity generation stage, they asked subjects to do some real effort tasks, and informed them of their absolute performance, and then told subjects that their final outcomes were actually randomly decided. It was not clear how people would identify themselves with two pieces of information. For example, if a person got a low score but a high payoff, would he build a bond with a low-payoff recipient for the reason that he thought he deserved low payoff, or with a high-payoff recipient simply for the same payoff? It is hard to identify the in-group bias with the ambiguous relationship between outcomes and identities. My experiment resolves this concern. In the identity generation stage, participants do the same task, and have the same performance. They spend several minutes experiencing the random generation process, so it is pretty clear that their payoffs are due to random luck only. Second, their experiment asked subjects to do redistribution decisions, which might weaken the effect of common experience. Shayo [2020] summarized that distance from the group attributes is one of the main factors that can influence the in-group bias. With the redistribution decision, decision makers change the outcomes of recipients, so that the common experience is only one element. My design keeps one's experience unchanged, and so it is easier to study how the common experience affects future decisions. Third, I have a control group where decision makers do the same task but are guaranteed a fixed payment. Unlike the Cassar and Klein [2019]'s design in which they used the subjects who experienced inequality in a tournament scheme as a control group, my control group provides a neutral benchmark without any self-serving bias, which might happen when the failures in the tournament thought results were more likely to be determined by luck than efforts and gave more to other failures for compensation.

This paper also contributes to the literature on the formation of group identity and in-group bias in the lab. Several experiments have shown that people would like to sacrifice their own benefits to increase in-group members' welfare [Chen and Li, 2009], cooperate with in-group members [Chen et al., 2014, Eckel and Grossman, 2005, Goette et al., 2006, McLeish and Oxoby, 2007], punish norm violations much more if the victim of the violation belongs to their group [Bernhard et al., 2006], and tell less truth to the out-group members [Rong et al., 2016].

However, there is only limited evidence regarding what generates group identity. Instead of priming participants' real identities (like Adnan et al. [2021], Benjamin et al. [2010], Bernhard et al. [2006], Chen et al. [2014], Goette et al. [2006], Hoff and Pandey [2006]), using the artificial group identities generated by the experimenter can help study the factors determines group affiliation. In most of the lab experiments with artificial groups, group identity was established by the group-solving task [Charness et al., 2014, Chen and Li, 2009, Eckel and Grossman, 2005, McLeish and Oxoby, 2007, Rong et al., 2016]. Some papers have also shown that a random assignment without any individual choices is sufficient to generate group affiliation [Chen and Li, 2009, Currarini and Mengel, 2016, Rong et al., 2016]. However, this finding has been challenged by other papers

[Charness et al., 2007, Eckel and Grossman, 2005]. Overall, limited methods have been used to generate group identities in the lab, and it remains under-studied what kind of activities can generate group affiliation. In my design, I plan to study whether common experience generates group affiliation, which is captured by distribution decisions. I show that people develop group affiliation asymmetrically. People who share the fortune with others perform more in-group favoritism by allocating more money to "in-group" members.

1.3 Experimental Design

In this section, I describe my experimental design for investigating the effect of common experience in distributional decisions. I have two treatments. The main treatment has subjects share common experience, and the control treatment has subjects not share common experience. In each treatment, I ask subjects to distribute extra money to a pair after an inequality generation process. I start with the main treatment, and address the difference between the control treatment and the main treatment next.

1.3.1 Main Treatment

In each treatment, subjects are asked to play two stages. In the original treatment, Stage 1 is to manipulate the common misfortune. Stage 2 is to elicit subjects' reparation preferences. The details of each Stage are demonstrated below.

1. **Stage 1 – Common Experience Manipulation**

   In Stage 1, subjects are asked to count the number of zeros in 10 tables as shown in Figure 1.1. They need to answer the number correctly to proceed, otherwise they need to count again for the same table. After finishing the counting task, they will learn their payments privately. Subjects' payments for Part 1 are randomly determined. 2/3 of them get \$3, and 1/3 of them get nothing. The randomization is implemented with the lottery number. I draw a lottery number for each of them. Who gets the number that is divisible by 3 gets \$0, others get \$3. This manipulation of payments allows me to separate subjects into two groups: the fortunate group, who gets what they deserve; the unfortunate group, who do the same task, but do not get what they deserve.

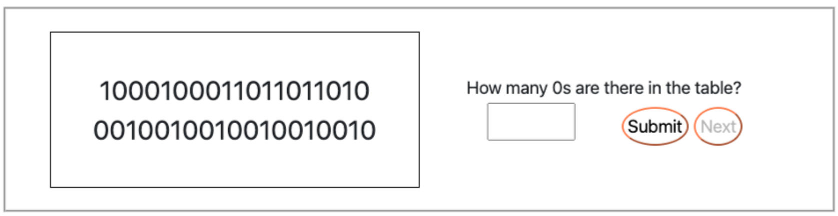


Figure 1.1: The Counting Task

2. **Stage 2 – Dis-intereted Allocator Game**

   In Stage 2, subjects are going to play a dis-interested allocator game. I use the strategy method [Brandts and Charness, 2011] to elicit allocators' strategy profiles. Everyone is going to allocate \$5 to the other two recipients, under three scenarios as shown in Figure 1.2: if both of them earn \$3 in Stage 1, if both of them earn \$0, if one of them earns \$3, and the other earns \$0. Every three subjects will be randomly grouped together to determine the payment for Stage 2. One of the three in a group will be assigned as an allocator, and she will get a flat payment \$$X$. In order to avoid self-comparison in distribution, X will be written on a paper, kept in view, upside down, and will be revealed at the end of Stage 2[1]. I set $X = 2$ in the main treatment. In total, I collect three choices from each subject.

[1] X has been used in Charness and Rabin [2002], Charness et al. [2008]

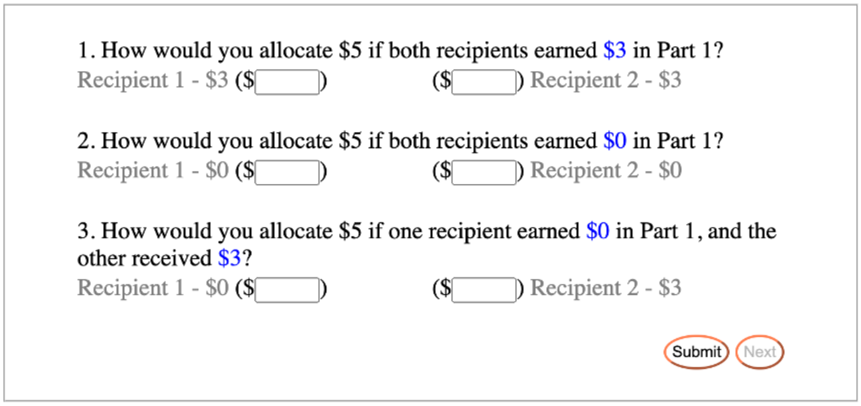


Figure 1.2: The Allocation Task

1.3.2 Control Treatment

In order to study the effect of common experience on distributional decisions, my control treatment needs to provide a benchmark of the decisions without common experience. Therefore, the difference between the main treatment and the control treatment measures the effect of interest.

The main feature of the control treatment is that dis-interested allocators share no common experience with any recipient, while they still allocate $5 to recipients in three different scenarios as in the main treatment. In order to satisfy this condition, I separate allocators and recipients in different sessions. In this setting, allocators still play two stages, while recipients do not make allocation decisions, and only play Stage 1. I collect distributional decisions from allocators only, and compare them with the decisions in the main treatment. I will demonstrate the details below.

Subjects as the control recipients only do Stage 1. They do the same counting task and receive the same random payments as in the main treatment. This part provides the same recipients for the allocators as the main treatment. They do not participant in Stage 2 and do not make any choices, so their payments are totally at the mercy of chance and control allocators.

Subjects as the control allocators do both Stage 1 and Stage 2, but they do them differently from the main treatment.

1. **Stage 1 – Experiencing Inequality Generation Task**

   In Stage 1 for the control allocators, subjects do the same counting task as in the main treatment. However, they receive a fixed payment of $\$Y$. I set $Y = 2$. This change allows the allocators to fully understand the procedure that generates inequality, while leaving them net of any common fortune or misfortune experience with recipients.

2. **Stage 2 – Dis-interested Allocator Game**

   In Stage 2 for control allocators, subjects will do the same allocation decisions as the main treatment. Subjects need to allocate \$5 to two control recipients, who might both receive \$0, or \$3 in Stage 1, or one of them receives \$0, and the other receives \$3. Subjects will receive $\$X$ for allocation completion. I set $\$X = 3$.

1.3.3 Payments

Each subject receives a \$5 show-up fee, and the payments earned in the experiment. In specific, the recipients in both treatments earn either \$3 or \$0 for Stage 1, and the money distributed from allocators for Stage 2. Allocators in the main treatment earn either \$3 or \$0 for Stage 1, and \$2 for Stage 2. Allocators in the control treatment earn \$2 for Stage 1, and \$3 for Stage 2.

1.4 Hypotheses

I assume there are two preferences when people make distribution decisions. When the allocator shares no common experiences with recipients, or the allocator has common experiences with both recipients, as discussed in the social preference literature, I assume people have preferences for inequality aversion in decision making. When people have preferences for inequality aversion, they dislike differences in payoffs, and try to equalize recipients' final payoffs. When the allocator shares common experiences with one recipient, I assume that common experiences establish a bond among people, and therefore the sense of group affiliation. As discussed in the group identity literature, when people have in-group bias, they want in-group members to receive more than out-group members. I hypothesize that subjects who received \$3 in Stage 1 would identify themselves as the fortunate group, and subjects who received \$0 in Stage 1 would identify themselves as the unfortunate group. The inequality-aversion preference and in-group bias together result in Hypothesis 1.

Neutral allocators tend to equalize recipients' final payoffs in all scenarios. Fortunate allocators and unfortunate allocators tend to equalize final payoffs when recipients are from the same group, but typically distribute more money to the in-group recipient if two recipients are from different groups.

In terms of in-group bias, I further assume asymmetry of magnitude participants distribute to the in-group members among different groups. Psychology literature has found evidence that negative events have a larger effect on people's behaviour than positive events (for example, Ito et al. [1998]), so I further assume that the effect of common misfortune is larger than common fortune, i.e. unfortunate allocators exhibit larger in-group bias. Some studies have also found that members of minority groups exhibit greater inter-group discrimination than members of majority groups (for example, Leonardelli and Brewer [2001]), I therefore assume that the effect of common misfortune is exaggerated in my experiment. The negativity bias and minority bias together result in Hypothesis 2.

Unfortunate allocators exhibit larger in-group bias than fortunate allocators, and de-

viate more from equalization than fortunate allocators when recipients are from different groups.

1.5 Results

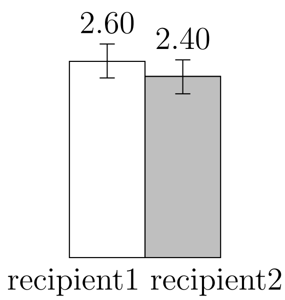
The experiment was run at the Experimental and Behavioral Economics Laboratory at UCSB, from February to April 2022. I collected 93 observations for the main treatment - 31 as the unfortunate allocator, and 62 as the fortunate allocator, and 25 observations for the control allocators. I also recruited 50 participants as control receivers, whose payments was decided by the control allocators. Subjects earned \$9.5 on average for 20-minute participation, with minimum \$5, and maximum \$13.

I present the results in three steps. First, I investigate whether people have preferences for inequality effect when recipients are from the same group. Second, I study whether the common experience of misfortune and fortune in the income generation stage affects distribution decisions. Third, I examine whether the common misfortune and common fortune have asymmetric effects.

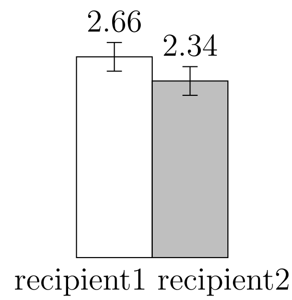
I first look at distributions to recipients from the same group. Figure 1.3 shows the distributions from different types of allocators. Panels on the left side show distributions to two unfortunate recipients who received \$0 in Stage 1, panels on the right side show distributions to two fortunate recipients who received \$3 in Stage 1. Unlike Hypothesis 1 that there is no difference between distributions to recipients from the same group from any type of allocators, we find an overall significant difference between a random recipient 1 and a random recipient 2 (two-tailed paired two samples Wilcoxon test p = 0.023). This is driven by the top right panel, distributions to two fortunate recipients from the fortunate allocators. I find that fortunate allocators significantly distribute 14% more to one recipient than the other, (two-tailed paired two samples Wilcoxon test p=0.085). There is no significant difference in distributions between recipients for other panels (two-tailed paired two samples Wilcoxon test p>0.2 for each case). Even though we observe different variance among different allocators, we found the majority (74%) of the distributions are consistent with Hypothesis 1 that there is no big difference between distributions to recipients from the same group. I summarize these findings in Result 1.

Result 1. *When distribute money to recipients from the same group, almost all allocators choose to equalize the distributions, except for the fortunate allocators, who on average distribute slightly more to one recipient than the other when they distribute money to two fortunate recipients.*

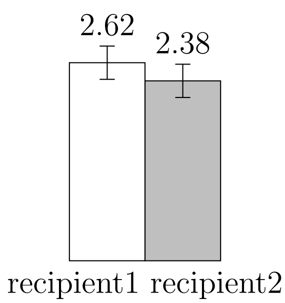
I then look at distributions to recipients from different groups. Figure 1.4 depicts the means of these decisions. The control group provides a benchmark net of common experience. If Hypothesis 1 is true, we would expect they distribute $4 to the unfortunate recipient and $1 for the fortunate recipient. Treatment groups capture the effects of common experiences in addition. The middle panel shows the distribution from neural allocators. They slightly deviate from equalization distributions, and leave more money to the fortunate recipients (two-tailed one sample Wilcoxon rank-sum p=0.029). The left panel shows that the fortunate allocators distribute on average about 40% ($0.6) more to the fortunate recipient than the neutral allocators (one-tailed unpaired two samples Wilcoxon rank-sum p=0.054). The right panel shows that the unfortunate allocators distribute on average about 10% ($0.35) more to the unfortunate recipient than the neutral allocators (one-tailed unpaired two samples Wilcoxon rank-sum p=0.089). The direction of the treatment effects are consistent with Hypothesis 1, but the decisions from neutral allocators are slightly different. However, it is too early to make a conclusion about the effect of common experiences because it is likely that the results are driven by some extreme cases. We next look at the distributions of these decisions.



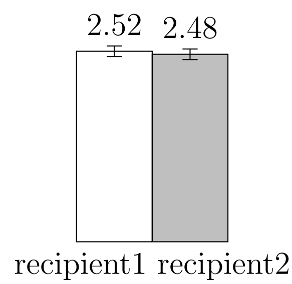
(a) fortunate allocators,
unfortunate recipients
N=62



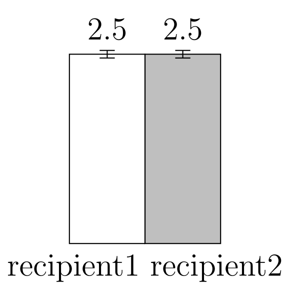
(b) fortunate allocators,
fortunate recipients
N=62



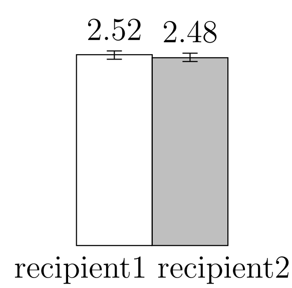
(c) neutral allocators,
unfortunate recipients
N=25



(d) neutral allocators,
fortunate recipients
N=25

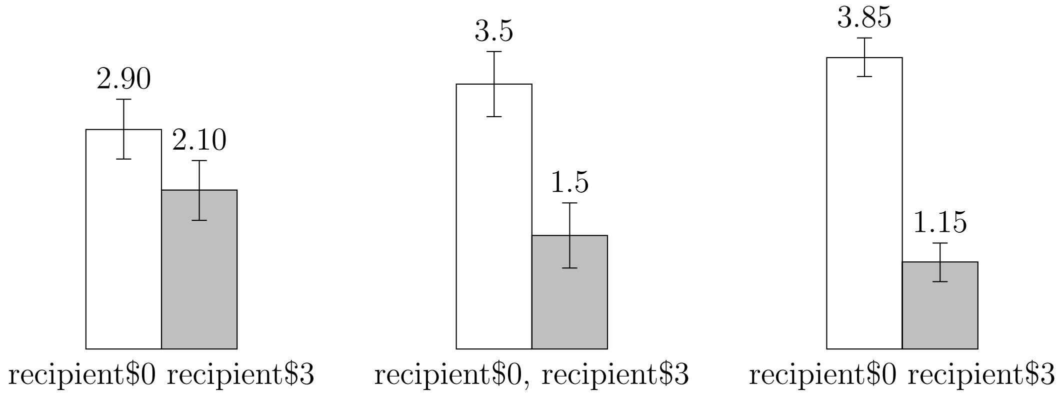


(e) unfortunate allocators,
unfortunate recipients
N=31



(f) unfortunate allocators,
fortunate recipients
N=31

Figure 1.3: Distributions to recipients from the same group. The heights of the bars and the values on top of the bars correspond to the means of the amount distributed to the recipient. Spikes at the top of the bars are equal to the 95% of confidence intervals of the means. Fortunate participates are those who received \$3 in Stage 1, and unfortunate participants are those who received \$0 in Stage 1.



(a) fortunate allocators N=62 (b) neutral allocators N=25 (c) unfortunate allocators N=31

Figure 1.4: Distributions to recipients from different groups. The heights of the bars and the values on top of the bars correspond to the means of the amount distributed to the recipients. Spikes at the top of the bars are equal to the 95% of confidence intervals of the means.

Figure 1.5 depicts the cumulative distribution functions (CDF) of decisions to the unfortunate recipients. There is no big difference if we compare the decisions from the fortunate allocators and neutral allocators, and from the unfortunate allocators and neutral allocators (one-tailed Kolmogorov-Smirnov test $p>0.30$ for two comparisons). This might be driven by the fact that the majority of allocators of each type tend to equalize recipients' final payments. In specific, 55% of fortunate allocators, 68% of neutral allocators and 74% of unfortunate allocators choose to do so. Even though the difference in percentage choosing equalization between each treatment group and the control group is mild, I notice the difference between two treatment groups is larger, which might lead to a significant difference in CDF. When I compare the decisions from the fortunate allocators and unfortunate allocators, the difference is significant (one-tailed Kolmogorov-Smirnov test $p=0.045$). This is partially in line with Hypothesis 1 that fortunate allocators distribute more to the fortunate recipients than unfortunate allocators. The CDF further suggests that the average difference is driven by a subset of fortunate allocators who distribute all money in favor of the fortunate recipients (approximately 16%), and the fact that no unfortunate allocator distributes more than \$3 to the fortunate recipients, which indicate that the in-group bias coming from the effect of common experience is at play. Hence, I conclude the following result.

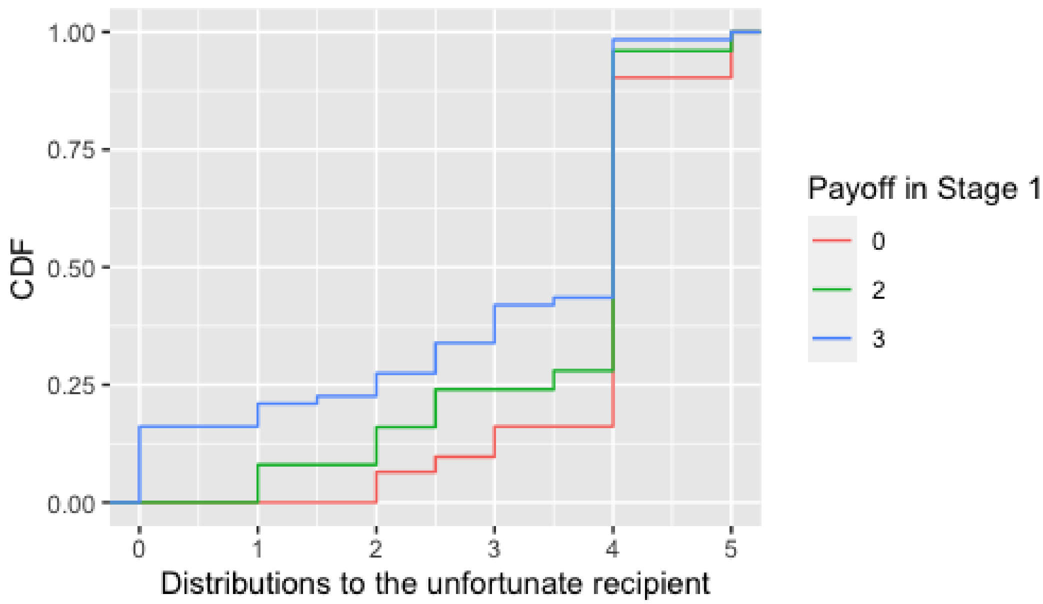


Figure 1.5: Cumulative distribution functions of decisions to recipients from different groups.

Result 2. *More than half of allocators choose to equalize recipients' final payoffs when they come from the same group in each treatment. Even though there is no big difference in decisions between each treatment and the control group, unfortunate allocators significantly distribute more money to unfortunate recipients than fortunate allocators do, and vice versa.*

Finally, I use the following econometric model to test the magnitude of the effect of common experiences:

$$Y_i = \alpha + \beta_1(UF)_i + \beta_2(F)_i + \epsilon_i, \tag{1.1}$$

In the model above, Y_i is how much the allocator distributes to the unfortunate recipient when recipients are from different groups; $(UF)_i$ equals one if the allocator is unfortunate; $(F)_i$ equals one if the allocator is fortunate; Therefore, α measures how much on average a neutral allocator distributed to an unfortunate recipient, β_1 measures how much more on average an unfortunate allocator distributed to an unfortunate recipient, β_2 measures how much less on average an fortunate allocator distributed to an unfortunate recipient.

Table 1.1 shows the results of regressions. We can see that only common fortune experience has a significant effect on allocation decisions, and they distribute 21% less to the unfortunate recipients than others. I then conclude result 3.

Result 3. *The effect of common misfortune experiences is mild, while the effect of common fortune experiences is significant.*

This observation strongly violates Hypothesis 2 that negative events have a larger impact than positive one, which means that pure in-group favoritism might not be the driving force for these decisions, and the common experience affects people's decisions through other ways. I came up with two alternative mechanisms for these results in Section 6.

1.6 Discussion

The main result is that there is no big difference between decisions from the neutral allocators and unfortunate allocators, and fortunate allocators significantly favor in-group members when two recipients are from different groups. This result strongly violates Hypothesis 2 and partially violates Hypothesis 1. In view of these results, I introduce two frameworks that might give some insights.

	Dependent variable: Distribution to the unfortunate recipient ($)
misfortune	0.355
	(0.342)
fortune	-0.597*
	(0.301)
Constant	3.500***
	(0.254)
Observations	118
R^2	0.098
Adjusted R^2	0.083
Residual Std. Error	1.272 (df = 115)
F Statistic	6.261*** (df = 2; 115)

Note: stardard error in the parenthesis, $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 1.1: OLS Regression

1.6.1 Alternative 1

People all have preferences for inequality aversion. In addition, people might want to legitimize their payoffs when they got above average for doing the same task. This preference echoes with Cherry et al. [2002], Oxoby and Spraggon [2008] who found that people are more likely to protect their and others' legitimate assets than illegitimate assets. Their results suggest that when people have a choice, they would like to legitimize their assets. Not penalizing others who got the same payoff as they did is a good signal that their payoffs are legitimate. Since neutral allocators did not get paid above average, and they shared no same payoff with any recipient, they did not need nor could legitimize their payoffs through distribution decisions. As a result of inequality aversion, they equalized final payoffs for all recipients. This is consistent with the result that equalization decisions accounts for 68% of all neutral allocators' decisions. Unfortunate allocators did not need to legitimize their payoffs since their payoffs were below average, and therefore they tried to make fair allocations as neutral allocators. This is consistent with the result that equalization decisions accounts for 74% of all unfortunate allocators' decisions. Fortunate allocators got above average, and they might want to legitimize their payoffs. The extreme case for not penalizing others who got the same payoff is to reward them. The highest reward is to allocate all \$5 to the fortunate recipient. This is consistent the second most popular decisions among fortunate allocators, and accounts for 16% of all fortune allocators' decisions. Not all people have the preference for legitimization when they get above average. If this is the case, the inequality aversion would be the driving force as for other people, and leads the allocators to choose equalization decisions. This is consistent with the most popular decisions for fortunate allocators, which accounts for 55% of their decisions. In total, this framework explains 71% of all decisions.

1.6.2 Alternative 2

People have preferences for inequality aversion and in-group favoritism. Since neutral allocators had no sense of group affiliation, their choices would be driven by inequality aversion. For other allocators, there was a tradeoff between fairness and in-group bias in their distribution decisions. I assume that deviation from fairness generates a fixed cost for participants. This echoes with Andreoni and Bernheim [2009], who found that people have a strong preference for the 50-50 norm. I further assume that all people care more about in-group members' payoffs than out-group members' payoffs, and the more the in-group member is ahead, the merrier they are. The preference function looks like below.

$$U(x) = \alpha(2x - 8) - \beta \mathbb{1}\{x \neq 4\} \tag{1.2}$$

In the function, \$$x$ is the final payment distributed to an in-group member. Since the sum of final payoffs for the two recipients is \$8 (\$3 in total for Stage 1, and \$5 in total for

Stage 2), $\$8-x$ is the final payment for the out-group member, and $\$2x-8$ is how much the in-group member is ahead. Therefore, $\alpha(2x-8)$ captures the utility from in-group bias. If the allocator equalizes final payoffs, he would leave \$4 to the in-group member, otherwise, he would suffer a fixed cost β of deviation from equalization.

Equalization gives the allocators utility of 0. The allocators deviate from equalization if and only if $\alpha(2x-8) > \beta$. Given α and β, the more the in-group members could be ahead, the more likely they would deviate from equalization. In my setting, the choice set of the final payoff to the in-group members is different for different allocators. For fortunate allocators, choosing between \$0 and \$5 for in-group members in Stage 2 means they can leave the in-group members with payoffs between \$3 and \$8. Therefore, they can let the in-group members be ahead by at most \$8. For unfortunate allocators, choosing between \$0 and \$5 for in-group members in Stage 2 means they can leave the in-group members with payoffs between \$0 and \$5. Therefore, they can let the in-group members be ahead by at most \$2. It's likely that the threshold of $2x-8$ lies in the interval between \$2 and \$8 for people who care about in-group favoritism. Therefore, Fortunate allocators had choices above the threshold while unfortunate had no such choices, and we only see in-group bias among fortunate allocators. Among fortunate allocators who did not equalize recipients' payoffs, 73% of them chose the inequality level above \$2. This might support this explanation.

1.6.3 Future Experiment to Distinguish Two Frameworks

Alternative 1 assumes people have different preferences, while Alternative 2 assumes people have the same preference, and it is the environment that blocks the manifestation. The way to distinguish between Alternative 1 and Alternative 2 is to uniform the environment for all the participants. For example, I can restrict the choice set to ask allocators to distribute \$7 and require allocators to distribute at least \$3 to the unfortunate recipients. Therefore, the in-group biased inequality choice set for all allocators is from \$0 to \$4, and the fortunate allocators still have a chance to reward the in-group member for distributing more than the half to the in-group members. With the change, Alternative 1 predicts no difference for the fortunate allocators and unfortunate allocators' decisions since fortunate ones still can make signals for legitimacy and unfortunate still only care about fairness. Alternative 2 predicts that more unfortunate allocators would choose to favor in-group members rather than fair allocations because deviation from fairness becomes more profitable for them.

Chapter 2

The Impact of Response Time on Deceptive Communication

Joint work with Bohan Ye

2.1 Introduction

Numerous economic interactions encompass a range of indirect choice information that can profoundly influence the outcomes of interactions. These supplementary elements may take the form of diverse data, signals, or characteristics that are not directly tied to the decisions being made but nevertheless offer valuable insights or contextual understanding. In today's technologically advanced environment, Response Time (RT) serves as a highly convenient metric for the decision-making process. The emergence of modern communication technologies, including Slack, emails, and instant messaging, has significantly facilitated the tracking of RT, making it more practical and accessible.

RT can potentially be a valuable tool in addressing asymmetric information in a sender-receiver game where there is a conflict of interest. In scenarios where a car dealer is driven by the constant desire to sell the most expensive cars, there is a possibility that they may affirmatively market the highest-price vehicle in your price range without adequately verifying if the car is your best match. Similarly, influencers seeking likes and followers may comment or re-post popular articles without genuinely engaging with the content or thoroughly reading it, a home inspector who wants to promote a deal may not meticulously examine each piece of technology before providing a report to a prospective consumer. In all these examples, senders have incentive to lie only in one direction, and uncovering the truth is time costly. It is reasonable to assume that RT may provide private information about the truthfulness of the message when there is an apparent conflict of interest, i.e., the faster the message, the less likely the sender uncovers the truth, the less credible it is. Van de Calseyde et al. [2014] have found that most of people trust the message with longer RT.

This research mainly aims to delve into what private information RT carries when there is a conflict of interest, and if there is a pattern between RT and truthfulness, tailored to a situation where truth discovery is time consuming. We're also intrigued by the strategic use of RT by both senders and receivers and understand how it contributes to their overall well-being. By unpacking the private information conveyed by RT and heuristics employed by receivers, we are able to make implications about how to improve the welfare of uninformed parties who used to be non-experts and non-authorities.

A cheap talk game models the sender-receiver interaction with a conflict of interest where a message is costless to transmit and receive. It is in contrast to signaling in which sending certain messages may be costly for the sender depending on the state of the world. In such a game, the sender gets private information about the truth and sends a message to a receiver. The message may disclose full information, partial information, or nothing at all. The receiver receivers the message, updates her belief of the truth and takes an action. The conflict of interest lies in that the sender always wants the receiver to take a certain action, like buying the most expensive car, while the receiver always wants to pin down the truth, like figuring out if the recommendation is the best match.

We have modified the cheap-talk game by introducing a discovery process at the outset of the game for the sender to invest effort in acquiring the truth. The difference from the conventional game is that now the sender does not exogenously get informed of the truth for free, instead, whether to become costly informed is an endogenous decision. In this adjusted setting, we define the RT as encompassing both the truth discovery process the deliberation process till the message sent.

We also modify the sender's utility by incorporating a truth telling preference. This preference is supported by the literature (see a survey conducted by Abeler et al. [2019]) that people forgo about 3/4 of the potential gain from lying in the individual game, and senders consistently transmit more information than theoretical predictions (see the summary in Lafky et al. [2022]).

In the modified game, we find that the sender would tell the truth if she discovers the truth. The intuition is that, if the sender would lie at a disadvantageous state after becoming informed, she'd better send the high-payoff message without becoming informed. In this way, she can avoid the cost of discovery while keeping the other expected payoffs the same.

If the sender behaves as the theory predicts, it is reasonable to assume that RT reveals if the sender uncovers the truth. The shorter the RT, the less likely the sender uncovers the truth, the less credible it is. If the receiver presumes this pattern, she would follow the long message more than the short message when there is an apparent conflict of interest. Given this obvious pattern, it's interesting to study if the sender would manipulate her RT if she knows that her RT will be observed by the receiver, and if the availability of RT will change sender's lying behavior.

We design an experiment to study the strategic use of RT. Specifically, we examine if the receiver uses RT effectively, and if the sender manipulates her RT when she is aware that it will be observed by the receiver. To answer the first question, we vary the availability of sender's RT in two treatments for the receiver: Receiver is Uninformed (RU) of sender's RT and Receiver is Informed (RI) of sender's RT. To answer the second question, we vary the awareness of RT in two treatments for the sender: Sender is Unaware (SU) of availability of RT and Sender is Aware (SA) of availability of RT. We ask each subject to play as a sender first, and then play as a receiver. In this way, all receivers have some experience about what RT may reveal before using it to make decisions.

Each subject plays three stages in an experiment. In Stage 1, subject plays as a sender and sends messages in 10 rounds. In each round, the sender decides whether to invest effort in learning the truth and then what message to send. We vary sender's treatment in Stage 1. The subjects either plays in the SU condition, or in the SA condition. In Stage 2, subject plays as a receiver receiving sender's messages only. This stage manipulates the RU condition. The receiver receives another subject's 10 messages from Stage 1 at once, and she takes 10 actions after each message. In Stage 3, subject plays as a receiver again, receiving the same sender's messages her RT. This stage manipulates the RI condition. The receiver receives the same 10 messages as in Stage 2 and corresponding RTs at once, and she takes 10 new actions after each message.

We recruited 62 subjects, 30 for the SU condition, and 32 for the SA condition. The most striking result is that senders uncovered the truth an impressive 92% of the time. More intriguingly, even when senders opted for a selfish approach, they still uncovered the truth 66% of the time. This results strongly violate our main theoretical prediction, and it implies that RT does not reveal if the sender uncovers the truth. The underlying reason might be that people are reluctant to lie, and they only lie if necessary.

Given such strong preference of getting informed of the truth, we have found that for the informed messages with a conflict of interest, the longer the RT, the higher probability of deceit. It suggests that lying is a hard decision for subjects, and requires more consideration. We also found that receivers were aware that RT might serve as an effective cue, and around half of them changed their decisions after observing RT at least once. However, they did not use it effectively. When receives had better trust more relative short messages, they did trust more long messages. We also found that senders were able to successfully manipulate their RT to their best interests, and availability of RT did not change their honesty rate.

These findings suggest that long RT is not an effective cue in detecting lies, which is also easily to be manipulated. Trusting toward the long RT is not an rational decision for receivers. Receivers might be better off if they just ignore the RT and trusting the message more when it's against the sender's interest. Speaking to the real life scenarios where it is hard to identify the discovery process, like when we ask the salesman or the physician for recommendation, if they spend more time than the time needed to figure out the truth, do not lean more trust to the recommendation with longer RT when there is apparent conflict of interest.

This research makes contributions to three areas of the literature. Firstly, it adds to the existing body of knowledge on the decision-making process by demonstrating that after the sender becoming informed, lying high-payoff messages take longer than truth-telling high-payoff messages, therefore providing evidence that serving self-interest requires a certain extent of deliberation. Secondly, it enriches the deception-detection literature by providing empirical evidence that receivers cannot effectively use the RT to detect lies, even though RT serves as a valuable cue. Thirdly, it advances the literature on the strategic use of RT by involving the information-seeking process with decision-making processes. Interestingly, in contrast with the moral-hazard predictions, most senders choose to uncover the truth even though they finally lie to the receivers. This result suggests that even for the sender who cares self-interest the most, she still hopes to not have to lie to get what she wants. So, she would uncover the truth if it is possible, and lies only if she knows for sure that she has to.

In the remaining sections of the paper, Section 2 provides a comprehensive literature review. Section 3 models the strategic lying game. Section 4 outlines the experimental design and hypotheses. Section 5 presents the results. Section 6 portrays a calibrated model. Finally, in Section 7, the paper concludes.

2.2 Literature Review

This research contributes to three strands of literature. The first strand focuses on the decision-making process of lying and truth-telling. The main focus of this literature is to identify the automatic tendency in decision-making, either serving self-interest or telling the truth. The dual-system approach [Kahneman, 2011, Rubinstein, 2016] would argue that the automatic tendency is always faster than the deliberate approach. To identify the genuine tendency without strategic interaction, the literature focuses on a paradigm introduced by Fischbacher and Föllmi-Heusi [2013]: subjects privately observe the outcome of a dice, report the outcome, and receive the payoff, either proportionally related to their report, or depending on the correctness of their guess. The key feature of this paradigm is that the experimenter does not observe an individual's truth, and therefore cannot identify any individual report as truthful or not. They can only identify decisions as high-payoff or low-payoff reports, and they distinguish people according to the difference between their observed reporting distribution and the statistically-predicted reporting distribution.

Greene and Paxton [2009] and Jiang [2013] have found that low-payoff reports take longer time than high-payoff reports. Greene and Paxton [2009] have also observed self-control neural activities (in the anterior cingulate cortex and ventrolateral prefrontal cortex) when dishonest people, whose high-payoff reports are much more frequent than predicted, forgo the gain. Shalvi et al. [2012] have found more high-payoff reports in the high-time-pressure condition than in the low-time-pressure condition. All the findings indicate that truth-telling requires extra self-control to resist the temptation, and is therefore a more deliberate approach, while serving the self-interest is the automatic tendency.

However, a limitation of this literature is the inability to distinguish between genuine high-payoff reports or deceptive high-payoff reports, therefore there is no solid evidence that self-interest is the automatic tendency. It's possible that self-interest also requires deliberation. Our design records the truth for each decision, therefore we are able to distinguish between the genuine high-payoff reports and deceptive high-payoff reports. We found that, contingent upon learning the truth, the latter takes longer than the former. This indicates that self-interest is not the automatic tendency, so there is some degree of hesitation. Moreover, our results show that when the transparency of RT is present, the same pattern remains even though senders condense the overall RT; this indicates that this pattern is difficult to manipulate.

The second strand of literature revolves around the identification of cues for detecting deception. This strand of literature centers on interactive games involving a minimum of two players with differing levels of information. This setup allows for the examination of scenarios where the sender has motivations to deceive and explores whether receivers can successfully discern veracity. Within this domain, researchers have explored two modes of deceptive communication: written messages and video clips.

In the case of written messages, receivers are able to detect lies by analyzing the content and linguistic characteristics. Charness and Dufwenberg [2006] found that written chat communication was highly effective in facilitating good social outcomes, with promises (and the subsequent changes in beliefs) being a key ingredient. Chen and Houser [2017] discovered that promises serve as reliable cues, leading receivers to place greater trust in them. However, other factors such as word usage and monetary references were found to be ineffective. On the other hand, video clips provide a more comprehensive set of potential cues, encompassing not only language but also nonverbal elements such as gender, facial expressions, body movements, and hand gestures. Studies conducted by Konrad et al. [2014], Dwenger and Lohse [2019], and Serra-Garcia and Gneezy [2021] revealed that receivers displayed only a slightly better than chance ability in detecting false reports of taxable income. In contrast, Belot and Van de Ven [2017] observed that receivers in the role of buyers were able to detect lies of sellers better than chance. Similarly, Bonnefon et al. [2013, 2017] demonstrated that trustors in a trust game exhibited limited ability to detect trustworthiness based on trustees' facial pictures. Todorov et al. [2015] also found that trustors easily formed first impressions from faces, although these impressions were unrelated to stable personality traits.

However, the existing literature on lie detection rarely addresses RT. Considering that previous studies, such as those conducted by Gneezy [2005] and Cai and Wang [2006], revealed that receivers tend to place greater trust in senders than what would be expected based on equilibrium predictions, there may be additional value in analyzing RT for receivers. Our research contributes to this domain by incorporating RT as a potential cue for deception detection. Our findings indicate that even when RT conveys private information, receivers fail to utilize it effectively. Furthermore, when RT is made public, receivers tend to exhibit an excessive degree of trust in senders, leading to slightly worse outcomes for the receivers.

The third body of literature focuses on the value of RT in strategic settings. This emerging literature discusses three main questions: when senders have private information, 1) whether their RTcarries private information; 2) whether receivers can extract private information from senders' RTs; 3) whether senders manipulate their RT. Frydman and Krajbich [2022] investigated the value of RT in a laboratory social learning game, while Konovalov and Krajbich [2023] examined the value of RT in a laboratory bargaining game. Both studies found that people's decision-making processes align with the drift-diffusion model [Fudenberg et al., 2018, Woodford, 2014], which states that the longer the RT, the smaller the difference between two choices. Cotet and Krajbich [2021] extended their results to the eBay market with experienced agents. Their findings implies that senders' RT carries private information, even in the field. In addition, Frydman and Krajbich [2022] and Konovalov and Krajbich [2023] demonstrated that receivers can infer private information from senders' RTs. Furthermore, Konovalov and Krajbich [2023] discovered that individuals attempted to manipulate their RTs when they were aware that their times would be revealed, making it less informative to their counterparts. This current research extends the investigation of the value of RT to the setting where it involves the information seeking process. The intuition is that given information acquisition is costly, senders should not uncover the truth when he decides to lie, and therefore lying should be very quick in such a setting, and very informative. However, our results show that many senders uncover the truth even if they decide to lie. In addition, when senders' RT carries private information, receivers are not able to take advantage of it. Finally, senders manipulate their RT to make it more informative. These unpredicted results indicate that there are more behavioral issues involved beyond preference of truth telling.

2.3 Theoretical Framework

In this section, we present our theoretical framework and discuss its main predictions. The model achieves two goals. First, it captures the relationship between the sender's RT and the type of the report. Second, it highlights the contrast between the situation when the sender is not aware of the availability of the RTand when the sender is motivated to manipulate it. These features generate a rich set of predictions that we then test experimentally.

2.3.1 Benchmark

In the benchmark, we study the perfect Bayesian equilibria for the cheap talk game with different preferences. We start from the standard cheap talk setting in which players only care about the monetary payoffs, and then extend the model to the situations where the senders have truth telling preference.

Standard Cheap Talk Game

Consider a game played between a sender and a receiver. There is a state of the world, drawn from a finite-state space. Both players do not care what specific state it is, instead, they care about if the state passes a certain threshold. In other word, there are only two payoff-relevant states. We assume that only payoff-relevant states matter for different strategies, therefore, it's equivalent to assume that there are only two states, $\Omega = \{\omega_L, \omega_H\}$. Both players have the common knowledge that the prior probability of the state ω_H is $\mu_0 \in (0, 1)$.

The sender observes ω and sends a message $m(\omega) \in M = \{\omega_L, \omega_H\}$ to the receiver. The receiver observes m (but not ω) and then chooses an action $a \in A = \{\omega_L, \omega_H\}$. For the message space, we lose some generality by limiting the number of message to two. Without allowing the sender to send nothing may force the sender to tell the truth more than the rate in some real life scenarios. However, the scope of this paper is not to estimate the truth telling rate, instead, we want to study if the RTcan help detect lies of a high-payoff message. Having more people telling the truth or lying provides us more useful data.

The receiver's payoff depends on ω, but the sender's payoff does not. We take $u_R : A \times \Omega \to R$ to be the receiver's utility, and $u_S : A \to R$ to be the sender's.

A strategy for the sender maps each state of the world to a distribution over messages $s_S : \Omega \to \Delta M$. A strategy for the receiver specifies a mixed action for her conditional on every message that she may observe $s_R : M \to \Delta A$. We are interested in studying the game's perfect Bayesian equilibria. The equilibirum consists of three elements, s_S, s_R, and a belief system $\beta : M \to \Delta\Omega$; such that:

1. the receiver knows the sender's strategy s_S, and, upon receiving the report m, updates her belief $\beta(\omega|m)$ regarding the state of the world using Bayes' law;

2. given belief β, s_R is optimal for the receiver, i.e., $a(m) = \arg\max_{a \in A} u_R(a, \beta(\omega|m))$, for each $m \in M$;

3. the sender knows the receiver's strategy s_R, and s_S is optimal given s_R, i.e., $m(\omega) = \arg\max_{m \in M} u_S(a(m))$, for each $\omega \in \Omega$.

We consider the setting that the receiver wishes to match her action to the state. That is, her state-dependent payoff is $u_R(a, \omega) = h$ if $a = \omega$, and $u_R(a, \omega) = l$ otherwise, where $h > l$. A rational receiver would choose $a = \omega_H$ when her posterior belief of the high-payoff state $\beta(\omega_H|\cdot)$ is larger than 50% and $a = \omega_L$ otherwise. We assume $\mu_0 < 50\%$. That is, with the prior belief, the receiver would choose $a = \omega_L$. The sender earns a higher payoff if the receiver chooses ω_H. Specifically, her payoff is $u_s(a) = h$ if $a = \omega_H$, and $u_s(a) = l$ otherwise.

In this setting, there are only babbling equilibria, in which the receiver infers no information at all from any message and sticks to the prior belief μ_0, and the sender sends random, uninformative message. The intuition is that if there is any message that the sender can send that will make the receiver choose ω_H, then in equilibrium the sender must send it. Hence the receiver will ignore the message.

Cheap Talk Game with Truth Telling Preference

According to a survey conducted by Abeler et al. [2019], people have a strong aversion to deception, and they forgo about 3/4 of the potential gain from lying in the individual game. This tendency persists in the strategic games. For strategic communication of private information, senders consistently transmit more information than theoretical predictions (see summary in Lafky et al. [2022]). These evidence suggests that people have truth telling preference. Theoretically, it's equivalent to add a cost of lying in the utility function.

Now, we modify the sender's utility by adding a fixed cost of lying c_θ. Formally, we assume her utility function

$$u_S(a, c(\omega, m; \theta)) = -z\mathbb{1}_{a=\omega_L} - c_\theta \mathbb{1}_{m \neq \omega}, z = h - l$$

As before, a is receiver's action, ω is the true state, m is sender's message. $c(\omega, m; \theta)$ denoting the individual cost of lying. θ captures the heterogeneity of subject's weighting that applies to the lying cost. When $c_\theta = 0$, the game is the standard cheap talk game. In this section, we focus on analysis with positive lying cost, i.e., $c_\theta > 0$.

To study the new equilibrium with the positive lying cost, we study the sender's behavior first. If the cost of lying is larger than the potential benefit of lying, i.e., $c_\theta \geq z$, truth telling is a (weakly when $c_\theta = z$) dominant strategy for the sender, so she would always tell the truth, i.e., $m(\omega) = \omega$ for each $\omega \in \Omega$.

If the cost of lying is smaller than the potential benefit of lying, i.e., $c_\theta < z$, sender's optimal strategy varies on the degree of the trust p from the receiver when the receiver follows the message with probability p and ignores the message and chooses $a = \omega_L$. It's worth noting that no matter how much trust the sender can get, it's always optimal for her to tell the truth at an advantageous state $\omega = \omega_H$. Because she needs to pay the lying cost and may suffer a monetary reduction if she lies downwards. Then we focus on what's optimal for the sender at a disadvantageous state $\omega = \omega_L$. If there is a great chance to gain the receiver's trust ($p > \frac{c_\theta}{z}$), the sender would lie upwards $m = \omega_H$ because the expected payoff gain covers the cost of lying. If the chance of trust is slim ($p < \frac{c_\theta}{z}$), the sender would tell the truth because the expected payoff gain is too small and is not worth it. If the degree of the receiver's trust is just enough to cover the cost lying ($p = \frac{c_\theta}{z}$), the sender may mix her message.

Another potential strategy for the low lying cost sender is to reverse her message, i.e., reporting $m = \omega_L$ for the high payoff h and reports $m = \omega_H$ for the low payoff l. The sender would lie downward in this strategy, i.e., report $m = \omega_L$ at the advantageous state $\omega = \omega_H$. This kind of strategy can be optimal only if the sender and the receiver shares the consensus that they should translate the message in the opposite way against its face value. We assume that the receiver would not do it, either because the face value of the message is too salient, or the receiver believes that there are enough honest people in the society. In our data, there are only 4 out of 620 choices that the sender lied downward. This result supports our assumption that players would communicate with the messages' face values.

We then turn to the receiver's strategy. Since the sender would never lie downward, the receiver should always follow the low-payoff message $m = \omega_L$. The receiver's optimal action given a high-payoff message $m = \omega_H$ depends on the likelihood q that the sender tells the truth at a disadvantageous state $\omega = \omega_L$, which is defined as $q = P(m = \omega_L | \theta = \omega_L)$. If the truth telling probability is huge enough ($q > \frac{1-2\mu_0}{1-\mu_0}$), it's more likely that the high-payoff message $m = \omega_H$ is true, and the receiver should trust the message $a = \omega_H$. If the likelihood is too small ($q < \frac{1-2\mu_0}{1-\mu_0}$), the receiver should ignore the message $a = \omega_L$. If the likelihood is just enough ($q = \frac{1-2\mu_0}{1-\mu_0}$) to make the receiver indifferent by taking different actions, she may mix her actions.

Which strategy profiles that are optimal to each other depends on the predetermined magnitude of the lying cost c_θ. If there is no incentive for the sender to lie at all, i.e., $c_\theta \geq z$, the sender would tell the truth and the receiver always follows the message. We call this equilibrium the truth telling one.

Otherwise, there is no pure strategy equilibrium. Suppose that the receiver always ignores the message and takes the optimal action $a = \omega_L$, the sender will tell the truth and get the low payoff l. It's worth noting that the sender will not lie because in that way she can get an even lower payoff $l - c_\theta$. Therefore, ignoring the message is not an optimal strategy for the receiver. So, there is no equilibrium with such strategy. Suppose that the receiver trusts the sender, the sender would always send high-payoff message $m = \omega_H$ for the high payoff. In this way, always trusting the sender is not an optimal strategy for the receiver. So, there is no equilibrium with such strategy either. There is a mixed-strategy equilibrium that the sender lies sometimes ($1 - q = \frac{\mu_0}{1-\mu_0}$) at the disadvantageous state $\omega = \omega_L$, and always tells the truth at the advantageous state $\omega = \omega_H$; the receiver always trusts the disadvantageous state message, and sometimes ($p = \frac{c_\theta}{z}$) trusts the advantageous state message. The intuition is that only if they confuses the other side, they would not be exploited.

With the positive lying cost, there can be a truth telling equilibrium, a partial lying equilibrium, and furthermore, there is no babbling equilibrium.

2.3.2 Cheap Talk with Truth Discovery Process

In the context of the lying cost model, we are introducing an additional step at the outset of the game. Rather than being automatically informed about the true state of the world, the sender now has the opportunity to make an informed decision about whether to invest time in uncovering the truth. This mirrors real-world scenarios where a salesperson makes a choice about whether to spend time checking which is the best match for the consumer, a home inspector needs to determine whether to meticulously examine each piece of technology before providing a report to a prospective second-hand buyer, or an individual must weigh the choice of approaching a director to inquire about job openings for a friend.

Let's begin our analysis by examining the equilibrium outcome, followed by an exploration of the connections between the sender's message and her RT. Additionally, we will investigate whether the sender's decisions would differ when she is conscious of the availability of RT.

Without RT

Consider a game that goes beyond the basic cheap talk game by adding an extra step. Now the sender is not automatically being informed of the truth. Instead, she faces a private decision d of whether to uncover the truth costly $d = 1$ by spending some time or not $d = 0$. Whether opting to become costly informed or remain uninformed about the truth, the sender subsequently conveys a message m to the receiver, and the receiver takes an action conditional on m. The key distinction for the receiver side, compared to the benchmark scenario, lies in the awareness that any message m could originate from either an informed state ω_L, an informed state ω_H, or an uninformed state.

We assume that discovering the truth incurs a cost, which could be associated with cognitive effort and time consumption, or keep ignorance can work as a moral wiggle room [Dana et al., 2007] that allows the sender to act more selfishly and provides extra utility. The sender's utility function changes to

$$u_S(f(d), a, c(\omega, m; \theta)) = -c_I \mathbb{1}_{d=1} + z\mathbb{1}_{a=\omega_L} - c_\theta \mathbb{1}_{m \neq \omega}.$$

Beyond the receiver's action a cost of lying $c(\omega, m; \theta)$, the sender's utility also depends on her discovering decision $f(d)$. She needs to pay a positive fixed cost c_I if she discovers the truth. Additionally, c_θ is predetermined prior to the revealing step.

If there is no lying cost at all, i.e., $c_\theta = 0$, the sender would babble if she has uncovered the truth. As a result, she would avoid the cost of discovering to keep uninformed and babble. Therefore, babbling is the unique equilibrium, and the sender would not uncover the truth at any time.

If sender's truth telling preference outweights the potential monetary gain and the cost of uncovering, i.e., $c_\theta > z + c_I$, uncovering the truth and delivering it is the dominant strategy for her. In this case, truth telling is the unique equilibrium and the sender would always get informed.

For the other positive lying cost, the receiver's strategy remains the similar since her information set does not change. With the no downward lying assumption, the receiver understands that the $m = \omega_L$ either from the informed state ω_L or uninformed state. Therefore, she should always trust such message, while the the action for the low-payoff message $m = \omega_H$ depends on the sender' truth telling rate q. The distinction from Section 3.1.2 lies in that the truth telling rate does not always from the known disadvantageous state, it might be the probability of discovering the truth.

The sender's strategy is different from the game with no truth seeking process. In this setting, the sender needs to consider whether she uncovers the truth first, and then what to do after each truth seeking decision. It's worth noting that the mixed the action after discovering the truth is a strictly dominated strategy. The sender can be strictly better off if she randomizes the action in the truth seeking decision than if she randomizes the truth telling rate after being informed of state ω_L. Specifically, instead of getting informed and lying sometimes (q) at the disadvantageous state, the sender can save some portion of the discovering cost by mixing the discovering decision: not discovering ($d = 0$) at the probability q and sending the high-payoff message $m = \omega_H$, and discovering ($d = 1$) the telling the truth ($m = \omega$) otherwise.

Proposition 1: The sender would report the truth if she chooses to be informed.

Given the discovering cost c_I, there are three strategies she has to consider for each trust rate p: 1) remaining uninformed and sending the low-payoff message ($d = 0, m = \omega_L$), 2) remaining uninformed and sending the high-payoff message ($d = 0, m = \omega_H$), and 3) getting informed and telling the truth ($d = 1, m = \omega$). The comparison between the first and the second strategies depends on the trust rate p. The comparison between the uninformed and informed strategies is related to the difference between the lying cost c_I and the cost of lying c_θ. When the discovering cost is too huge and the trust rate is too small ($c_I > \mu_0(pz + c_\theta), p < \frac{(1-2\mu_0)c_\theta}{z}$), the sender's optimal strategy to keep uninformed and report ω_L all the time. When the discovering cost is too huge or cost of lying is too small, and the trust rate is big enough ($c_I > (1 - \mu_0)(c_\theta - pz), p > \frac{(1-2\mu_0)c_\theta}{z}$), the sender's optimal strategy is to keep uninformed and report ω_H. For the rest of the cases, it's optimal for the sender to uncover the truth and report it. The sender would mix her strategies if there is no difference between one uninformed strategy and one informed strategy.

With positive lying cost, there is truth telling equilibrium as before when the truth telling preference is pretty strong. When the discovering cost is too huge to support the truth telling preference ($c_I > \mu_0 \times c_\theta$), there is babbling equilibrium, in the way that always reporting low-payoff message $m = \omega_L$. But there is no babbling equilibrium where the sender always reporting high-payoff message $m = \omega_H$. Because in this way, the sender cannot gain the receiver's trust and therefore she'd better report $m = \omega_L$ to save a portion of the lying cost. There is no equilibrium when the sender mixes the uninformed low-payoff message and the informed one. Because, in that way, the receiver should always trust the sender, therefore the sender is better-off if she remains uninformed and reports the high-payoff message all the time. There is a mixed strategy between an informed strategy and the uninformed and high-payoff one. The sender would discover the truth with probability $q = \frac{1-2\mu_0}{1-\mu_0}$ and tells the truth, otherwise sends the high-payoff message $m = \omega_H$. The receiver would always trusts the disadvantageous state message, and trust the advantageous state message with probability $p = \frac{c_\theta - \frac{c_I}{1-\mu_0}}{z}$.

In summary, there are three different types of babbling with positive discovering cost: getting uninformed and babbling, getting informed and truth telling, and partial discovering. The key distinction from the previous model is that the sender would always tell the truth if she gets informed of the truth, and in equilibrium, the suspicious behavior is always pertaining to keeping uninformed.

With RT

Now consider the game that the receiver possesses an additional piece of information about the sender's decision – her RT. This RT indicates the duration between when the sender initiates the process of making informed decision and when she eventually transmits her message.

Let's initially delve into the sender's behavior. In SU condition, the sender is unaware of the availability of RT, and there is no motivation for her to manipulate RT, and as a result, RT accurately reflects the genuine mechanical and decision-making process. If the sender does not consistently opt for being uninformed and babbling or consistently opt for being informed and telling the truth, RT contains private information about her type of message. Given *Proposition 1*, a very short RT suggests that the sender does not uncover the truth. Furthermore, if the RT carries private information, it's beneficial for the sender to manipulate the disadvantageous RT to be as long as the advantageous RT. The findings in Konovalov and Krajbich [2023] support this RT manipulation idea. We hypothesize that this relationship is weaker (but does not vanish) in the SA condition.

Hypothesis 1: In both SU and SA, the faster the y report, the less credible it is.

We then consider how the sender manipulates RTs in the SA condition, focusing on the sender who takes the mixed strategy. If the sender thinks the receiver uses the RT to detect lies, she may deliberately prolong her RT for an uninformed decision, feigning an extended process of uncovering the truth. This is a great opportunity for the sender to persuade the receiver that she tells the truth all the time and deserves the trust. In this way, the sender manipulates the RT gap between two types of high-payoff messages to be smaller in the SA condition than in the SU condition.

Hypothesis 2: In the SA condition, the sender will prolong her RT on average.

In the SA condition, we further assume that the method employed by the sender to extend her RT to a reasonable duration involves following the steps necessary to genuinely uncover the truth. If this is true, even the sender decides to serve the self-interest, she may uncover the truth in this scenario. Because the cost associated with lying after discovering the truth, at state ω_L, is greater than the expected value of the lying cost when remaining uninformed and sends the high-payoff message ω_H, represented as $c_\theta > \mathbb{E}c_\theta = (1-\mu_0)c_\theta$, the sender may be inclined to shift from their initial strategy of remaining uninformed to becoming informed and truthfully reporting. Consequently, in a situation where the sender is aware of the availability of RT, there is a general tendency for her to communicate more truth.

Hypothesis 3: In the SA condition, the sender is more likely to get informed, and are more likely to tell the truth than in the SU condition.

We then make hypotheses for the receiver's actions. If Hypothesis 1 is correct, the receiver can leverage RT to make more informed decisions instead of randomizing whether trust or not given the disadvantageous state message. When the sender's RT is exceedingly short, that is, when RT is significantly smaller than a certain minimum time threshold denoted as $\underline{t}$, it indicates that the sender did not invest enough time to uncover the truth. In this case, the receiver should make decisions based on the prior distribution and opt for ω_L. Conversely, when RT exceeds this minimum threshold, it implies that the sender may have taken the necessary time to discover the truth and is likely to report honestly. Consequently, the receiver should place more trust in the sender's message and act accordingly. Hypothesis 4a is based on the assumption that the receiver correctly presumes the pattern in Hypothesis 1. If it is true, the receiver is better-off by knowing RT, and is more better-off in the SU condition than in the SA condition.

Hypothesis 4a: The receiver is less likely to follow the short y report in the RI condition than in the RU condition. The difference between the two conditions is smaller in the SA condition than in the SU condition.

Another possibility is that the receiver may think that truth telling is automatic, and views the long RT as more suspicious, as the drift-diffusion model predicts, and therefore trusts the shorter y report more. If it is true, the receiver is worse-off by knowing RT, and is more worse-off in the SU condition than in the SA condition.

Hypothesis 4b: The receiver is less likely to follow the long y report in the RI condition than in the RU condition. The difference between two conditions is smaller in the SA condition than in the SU condition.

Last but not least, we're interested in the demographic characteristics in this setting. Gender is a salient demographic feature. The meta-analysis over 380 experiments conducted by Gerlach et al. [2019] suggests that men behaved slightly more dishonestly than women did. Hypothesis 5a extends the gender differences into truth uncovering, lying detection and manipulation of RT.

Hypothesis 5a: Females and males are different in uncovering the true states, truth telling, manipulating the RT, and detecting lies.

Creativity is considered one of the most important skills nowadays. A meta-analysis over 36 studies conducted by Storme et al. [2021] have revealed a weak positive correlation between creativity (measured via the self-report Gough scale), and dishonesty. Given that creative people are more likely to lie, we further conjecture that they are better at detecting lies. Hypothesis 5b posits such heterogeneity.

Hypothesis 5b: The subject who is creative is more likely to lie as a sender and is better at detecting lies as a receiver.

2.4 Experimental Design

In this section, we describe the laboratory implementation of our model, the main treatments that we conducted.

We begin by describing the implementation of the base game. Six dice are available to be rolled. For each die, there are six possible outcomes, 1, 2, 3, 4, 5 or 6. We define 4, 5, 6 as "large numbers", and 1, 2, 3 as "small numbers". The true states related to the outcome of six dice about whether there are 4 or more large numbers, yes (where there are 4 or more "large numbers" on 6 six-sided dice) or no (where there are 3 or fewer "large numbers"). The sender has a chance to uncover the outcome of each die. Her message can be yes or no.

The receiver's decision, along with the true state, determined the payoffs for both the sender and the receiver as listed in Table 2.1. The receiver earns \$8 if she correctly guesses the answer of the question. She earns \$4 otherwise. The sender earns \$8 if the receiver guesses that the answer is yes, irrespective of truth. She earns \$4 otherwise. Given this, the prior is $\mu_0 = 34.37\%$. To present our results, we adopt the following notation to distinguish between states, messages, and actions: $\omega = \{yes, no\}$, $m = \{y, n\}$, $a = \{YES, NO\}$.

Truth/Receiver's guess	YES	NO
yes	(\$8, \$8)	(\$4, \$4)
no	(\$8, \$4)	(\$4, \$8)

Table 2.1: Monetary Payoff: the first item denotes the sender's payoff, the the second item denotes the receiver's payoff.

We vary treatments in two dimensions as in Table 2.2. The first dimension revolves around the SU and SA conditions. We do not mention anything about RT to senders in the SU condition, and tell senders that their RT corresponding to each message will be recorded and provided to receivers in the SA condition. Each sender would only participant in one treatment, either SU or SA. So, we are able to make between-subject comparison for senders' behavior. The second dimension pertains to the RU and RI conditions. Receivers only receiver the information about senders' messages in the RU condition, and they receive additional information about senders' RT corresponding to each message in the RI condition. Each receiver would take part in two conditions, both RU and RI. So, we are able to make within-subject comparison for receivers' behavior.

	Sender Unaware (SU)	Sender Aware (SA)
Receiver uninformed (RU)		
Receiver informed (RI)		

Table 2.2: Treatments

We asked each subject to play as a sender first, and then play as a receiver. This will aid the receiver in acquiring insights into the expected RT of each round and the probability associated with the occurrence of advantageous states. There are three stages in the experiment as described in Figure 2.1. In stage 1, subject played the role of a sender, and they were randomly assigned to either the SU or SA condition. In the SU condition, senders were unaware if their RT would be provided to the receiver in the later stage. In the SA condition, senders were aware that their RT would be reported to the receivers.The sender faced an independent series of true states for 10 rounds. At the beginning of each round, 6 computer-generated dice was initially covered, and the sender decided if she would uncover each of them by clicking on the corresponding button. To find out the true state, senders would have to click on buttons to uncover at least three dice, which takes time. In fact, for all informed messages, senders clicked all 6 buttons more than 90% of times. The minimum of the time use for the informed message was 5 seconds in the SU condition, and 4 seconds in the SA condition. Senders have the freedom to uncover any number of the dice, including none, before sending a message. At the end of each round, the sender need to select a message, y (There ARE 4 or more large numbers of the 6 dice) or n (There ARE NOT 4 or more large numbers of the 6 dice), sending to the receiver. In total, each sender sent 10 independent messages at the end of stage 1.

In stage 2, subject played the role of a receiver. Each receiver would receive 10 messages from a non-self sender and had to guess the true state for each round, YES or NO. All 10 rounds of messages were provided at once to the receiver. This allows for the comparison of RTs across rounds in the RI condition, aiding in the assessment of whether a specific RT is fast or slow. In stage 3, subject played the role of a receiver again. In this stage, the receiver not only got the messages from the sender, also the sender's RT for each round. As in stage 2, all 10 messages and the corresponding RTs were provided at once.

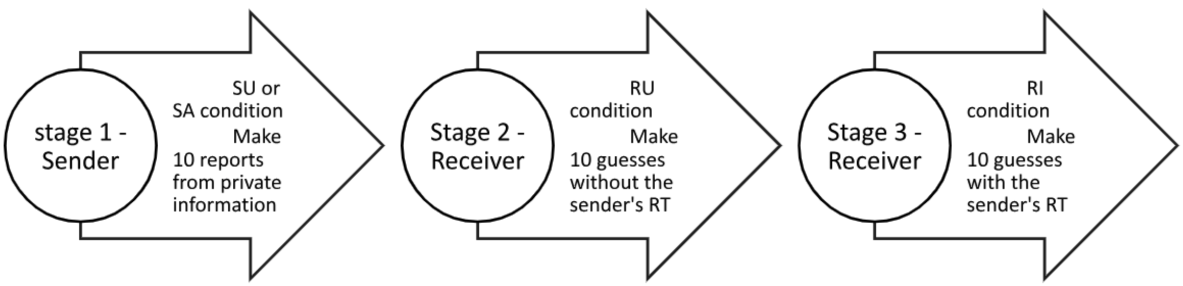


Figure 2.1: Experiment Timeline

Comprehensive questions were included in each stage to ensure that subjects understood the instructions. After completing the three stages, a survey was conducted to collect gender and personality information from each subject.

Each subject participated in 10 matches in RU condition, and 10 matches in RI condition as a sender and as a receiver. We paid them for two roles. To determine the payment, we rolled a 10-sided die twice. The first roll determined which condition would count: an odd number indicated that the RU condition would count, while an even number indicated that the RI condition would count. The second roll determined which of the 10 rounds for each role would count. For example, 2 and 9 mean that the subject would be paid for the 9th match as a sender in the RI condition, and for the 9th match as a receiver in the RI condition. Subjects were also paid an additional $1 for completing the survey.

The experiment was conducted at the Experimental and Behavioral Economics Laboratory at University of California, Santa Barbara, from May 11 to May 16, 2023. We recruited 62 subjects, 30 for the SU condition, and 32 for the SA condition. The subject earned an average of $13 for about 1 hour in the lab.

We collected 10 observations from each subject as a sender, and 20 observations as a receiver. In total, for the sender side, we collected 300 observations in the SU condition and 320 observations in the SA conditions. For the receiver side, we collected 620 observations in both the RU and the RI condition.

2.5 Results

We begin by examining hypotheses related to the behavior of senders. We then investigate hypotheses related to the behavior of receivers. We include heterogeneity tests of our hypotheses in the last part. The main finding is that while the longer response time indicates less credible report, we do find that receivers over-trust the delayed report. Overall, we conclude that the response time could be used as a cue in the deceptive communication when senders lack an incentive to manipulate, but receivers do not effectively leverage this cue.

2.5.1 Senders' Behavior

First, we test if senders' behavior is consistent with the theoretical prediction. Predicted by *Proposition 1*, when there is a cost of truth discovering, a rational sender would be truthful if she uncovers the truth, and all selfish decisions are made without being informed. We calculate the honesty rate of the informed message to test *Proposition 1*. Second, we examine if the relationship between RT and veracity of the report is consistent with *Hypothesis 1*, the longer the RT, the less credible of the informed report. Specifically, we investigate whether long y is more credible than the short y. Last, we test *Hypothesis 2* and *Hypothesis 3*, the difference between the SU and SA conditions.

To get an over view of the relationship between the RT and the type of the message, we calculated the average RT for each type of the report. We define the genuine message as the one that the sender uncovered enough buttons to get informed of the true state and reported truthfully, the deceptive report as the one that the sender got informed of the state and lied, the uninformed report as the one that the sender sent without uncovering enough buttons to get informed of the truth. Table 2.3 presents the number of each type of the report, and the corresponding average and standard error of RT.

SU (obs.)	Mean (s.e.)	SA (obs.)	Mean (s.e.)
genuine y (86)	12.10s (1.06s)	genuine y (87)	10.85s (0.74s)
genuine n (138)	13.48s (0.87s)	genuine n (160)	12.62s (0.66s)
deceptive y (51)	13.10s (0.95s)	deceptive y (45)	12.67s (0.83s)
deceptive n (1)	9s (0s)	deceptive n (3)	13.33s (1.33s)
uninformed y (24)	8.83s (2.34s)	uninformed y (25)	6.84s (1.27s)

Table 2.3: Mean and standard error of RT for difference messages

The ratio of genuine messages among all informed communications, representing the honesty rates of informed messages, are 81% and 84% under the SU and SA conditions, respectively. Specifically, the honesty rates for y reports are 63% and 66%, while for n reports, they are 99% and 98%. These findings indicate a significant deviation from

Proposition 1, which expects 100% honesty rates among informed reports. This suggests that either the cost of discovering the truth is not substantial enough, or other preferences are influencing senders' decisions, challenging of using the clear-cut RT to detect lies.

The low cost of discovering the truth typically predicts a low rate of discovery. Dana et al. [2007] and Grossman [2014] demonstrate that individuals tend to not actively seek out the truth, instead exploiting moral wiggle room to act in self-interest in the dictator game. Contrary to these findings, our study reveals that senders uncover the truth 92% of the time in both the SU and SA conditions, significantly higher than the 55% reported in previous research. Among the senders who uncover the truth at least once, 90% always do so (93% in the SU condition, 81% in the SA condition). This strong tendency to uncover the truth suggests that other preferences are at play in the cheap talk game, distinguishing it from the dynamics observed in the dictator game.

The key difference between the cheap talk game and the dictator game lies in the sender's role. In the cheap talk game, the sender influences the final decision by shaping the receiver's perception of the truth, whereas in the dictator game, the solo decision-maker directly chooses the outcome without affecting others. People's demand for information may vary across different contexts. While they may be reluctant to learn about potential harms to others, they exhibit a willingness to be informed in situations where lying is necessary.

One plausible explanation for the heightened demand for truth in our experiment is the preference for honesty in achieving objectives. Unlike the truth-telling preference, which pertains to an individual's inclination to be truthful once aware of the facts, this preference signifies a proactive pursuit of honesty when it serves one's goals. Senders opt for honesty unless lying becomes essential—only if they believe honesty cannot achieve their objectives. In our experiment, fabricating report y while remaining uninformed is unnecessary; achieving the same payoff through honesty is feasible. Lying becomes imperative only when the sender discerns the disadvantageous state $\omega_L = no$ and endeavors to exploit the receiver's trust for a higher payoff. This necessity arises only after the sender becomes informed of the truth. Thus, senders persist in uncovering the truth irrespective of incentives.

Given the violation of *Proposition 1*, RT seldom reveals whether the sender uncovered the truth. Here comes our first main result.

Result 1: RT in the cheap talk game with truth discovery process does not reveal if the sender uncovers the truth.

In addition to the time taken to discover the truth, RT also includes the deliberation process regarding which message to report in each round. Given that senders uncover the truth 92% of the time, RT primarily reflects the time spent in deliberation. Investigating whether RT correlates with the veracity of the report essentially examines whether there is a difference in the deliberation process based on varying incentives. According to existing literature, selfish decisions tend to be automatic [Greene and Paxton, 2009, Jiang, 2013, Shalvi et al., 2012]. Therefore, we would expect the RT for the n report to be longer than for the y report, as senders rarely lie from y to n,[1] and the y report often involves deception when senders lie from n to y as stated in Table 2.3. This expectation is confirmed: the average RT for the n report is 13 seconds, 2 seconds longer than the mean RT for the y report. This pattern holds true in both the SU and SA conditions[2]. However, the difference in RT between the different reports does not provide additional value beyond the face value of the report. RT can be value added if it is correlated with

[1]There is no uninformed n in both conditions. The honesty rates of the informed n are 99% in the SU condition and 98% in the SA condition.

[2]According to Ordinary Least Squares regression, the difference is not significant in SU condition, while it is significant in SA condition.

the veracity among y reports.

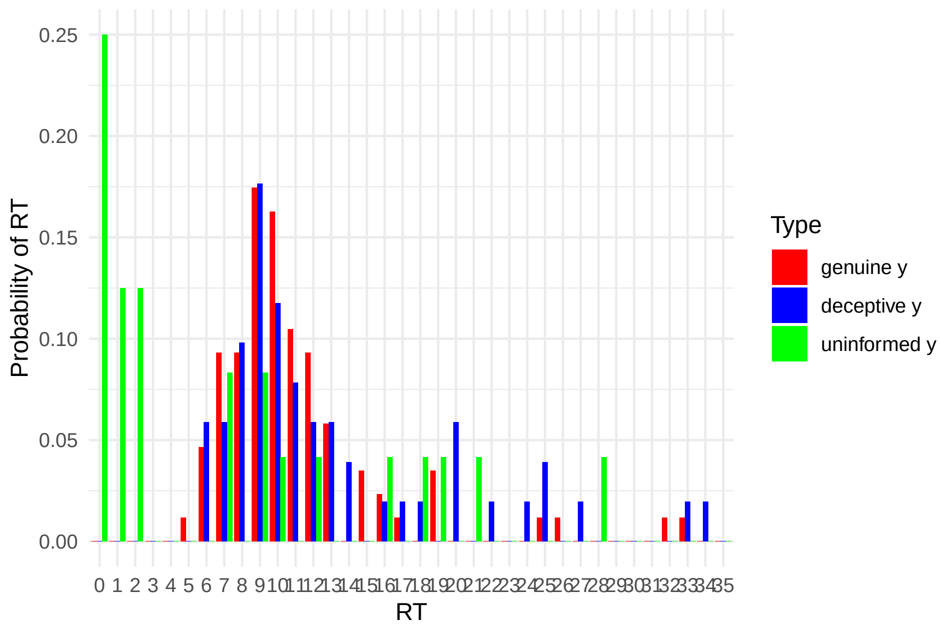
Given the overall 46% dishonesty rate (including deceptive and uninformed messages) for y reports, RT may work as an effective cue for detecting deceptive and uninformed y reports. This potential has been understudied in the literature. To address this, we investigate the relationship between RT and the veracity of y report in the SU condition, where the senders had no incentive to manipulate RT. We then examine if this relationship is apparent to the receivers. Finally, we study the relationship in the SA condition to understand if senders manipulate RT when they are aware of its availability.

Table 2.3 reveals that, on average, the RT of uninformed y is shorter than that of genuine y, while the RT of the deceptive y is longer than genuine y. This observation implies that lying to achieve a higher payoff entails a cost, manifesting in the longer time required compared to being honest to attain the same outcome. It also suggests that RT can be used to detect lies if the variance in RT for different types of y reports is not too large.

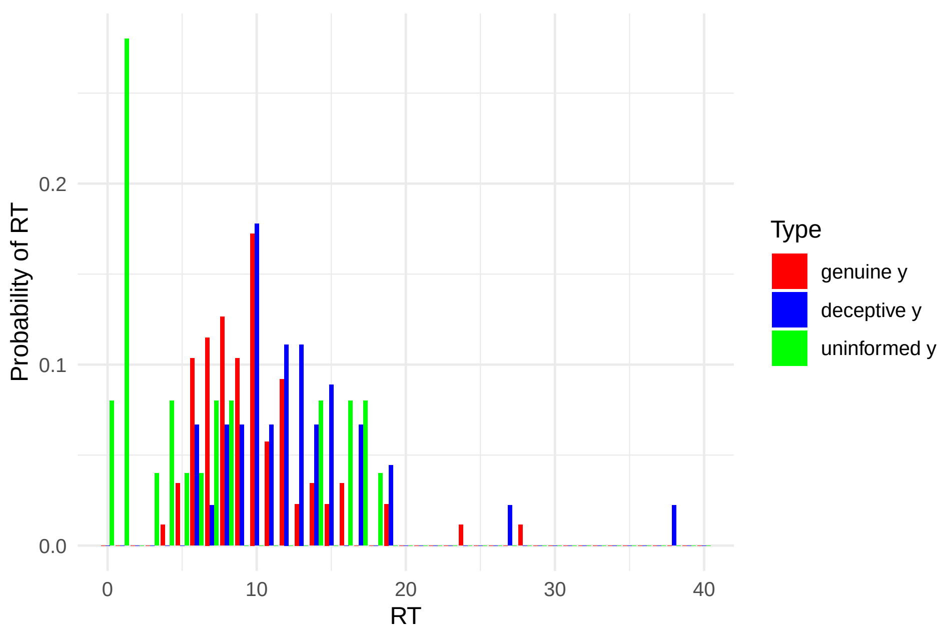
Figure 2.2a illustrates the distribution of RT for each type of y reports. Given that the vast majority of the RTs are within 35 seconds, we exclude two rare and extreme points with RTs of 47 and 90 seconds for simplicity.[3] The green bar represent the RT distribution for the uninformed y, the red bars for genuine y, and the blue bars for deceptive y. The presence of only green bars on the left end of the RT scale indicates that all immediate reports are uninformed. The predominance of blue bars over red bars in the right half of the RT scale suggests that reports with longer RTs are more likely to be deceptive than genuine. These results suggest that RT can be used to detect the genuine y reports effectively.

However, as not every sender engages in both deceptive and honest behavior across

[3]The message with RT=47 represents an uninformed y, while the report with RT=90 represents a genuine y.



(a) RT Distribution in SU. We exclude two rare and extreme points with RT=47 (uninformed y) and RT=90 (genuine y).

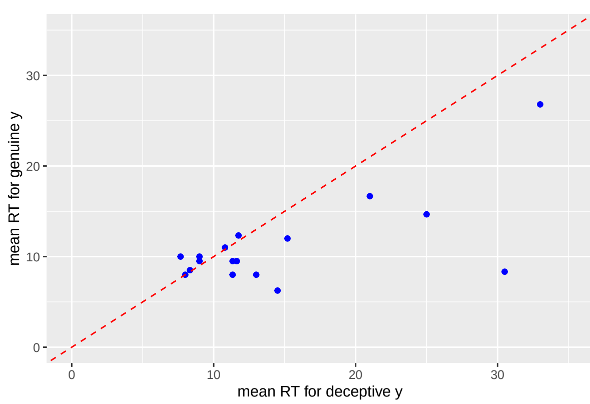


(b) RT Distribution in SA. We exclude two rare and extreme points with RT=44 (genuine y) and RT=51 (genuine y).

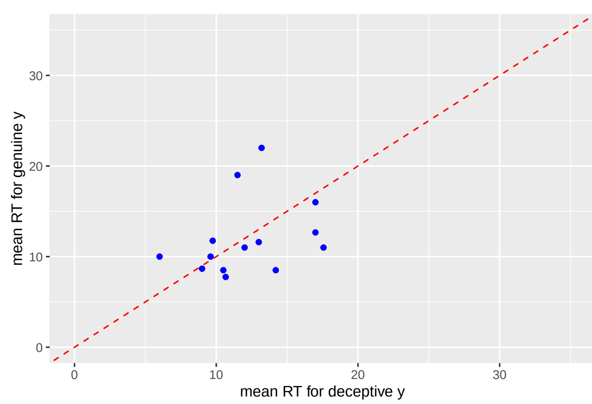
Figure 2.2: RT Distribution for Three Types of y Reports.

all 10 rounds, the mean and the distribution of the overall RT may lack consideration of individual differences. Indeed, in the SU condition, 11 out of 30 subjects tell the truth all the time, while 17 of them report both genuine and deceptive y, and the remaining 2 subjects always babble without being informed. As each receiver exclusively engage with a single sender, it becomes imperative to explore potential differentials in RT at the individual level.

Figure 2.3a illustrates the relationship between the RT for genuine y and deceptive y for senders who engage in both truthful and deceptive reporting over 10 rounds in the SU condition. Each data point represents the mean RTs of an individual sender. The x-axis denotes the average RT for deceptive y reports, and the y-axis denotes her average RT for genuine y reports. Point falling on the 45-degree line indicate no difference in RT between these two types of reports. The figure represents data from 17 subjects in the SU condition. The majority of data points are located in the lower triangle, implying that, for the same sender, deceptive y reports take longer time than genuine y reports.



(a) SU condition



(b) SA condition

Figure 2.3: Relationships between RT for genuine and deceptive y. Each data point represents an individual sender. The x-axis denotes the average RT for deceptive y report(s), and y-axis denotes the average RT for genuine y report(s). The red dashed line is the 45-degree line.

To test significance of the difference, we employ a fixed effect model as equation 2.1.

This analytical approach effectively controls for unobserved sender-specific factors that may drive the differences in means, thereby allowing us to robustly assess the nuanced disparities in RT across various report types. In addition to addressing the concerns, we also account for the number of rounds in our analysis. This adjustment is necessary as decision-making speed tends to increase over time due to the heightened familiarity with the context.

$$RT_{ir} = a_i + b \times Report_{ir} + d \times r + \epsilon_{ir}. \tag{2.1}$$

Within our analytical framework, RT_{ir} is sender i's RT at round r. $Report_{ir}$ assumes a crucial role by effectively categorizing sender i's report at round r into five distinct classifications: genuine y, deceptive y, genuine n, deceptive n, or uninformed y. We take the genuine y as the benchmark, and b capture the differences between any other report and the genuine y. We also control the time trend (see Appendix B.1) through incorporating the round number and control individual idiosyncrasies through a_i. ϵ_{ir} captures the random noise.

Table 2.4 describes the ordinary least squares regression results. Column 2 presents the outcomes conducted under the SU condition. Notably, within the realm of y reports, deceptive ones take significantly longer—by 4 seconds—compared to genuine ones ($p < 0.03$). In contrast, uninformed y reports take approximately 5 seconds less, though this difference lacks statistical significance ($p = 0.35$) due to the relatively infrequent occurrence of such behavior among senders. These results remain robust when excluding round 1, where subjects need time to familiarize themselves with the task, or when analyzing only the last 5 rounds, by which time subjects are well-acquainted with the task. These findings imply that, aside from extremely short RTs, a longer RT for a y report is associated with a higher probability of deception in the SU condition, and RT might serve as a cue to detect lies if the probability of a long y report being deceptive is sufficiently high.

	Dependent variable: RT	
	SU	SA
Genuine n	1.733	2.028**
	(1.197)	(0.874)
Deceptive y	3.561**	1.167
	(1.632)	(1.275)
Deceptive n	-5.615	0.109
	(8.788)	(3.858)
Uninformed y	-4.623	−4.093**
	(4.909)	(2.074)
Round r	−1.176***	−1.145***
	(0.167)	(0.117)
Constant	17.680***	15.665***
	(2.931)	(2.085)
Individual Fixed Effect	✓	✓
Observations	300	320
R^2	0.358	0.445
Adjusted R^2	0.276	0.374
Note:	*p<0.1; **p<0.05; ***p<0.01	

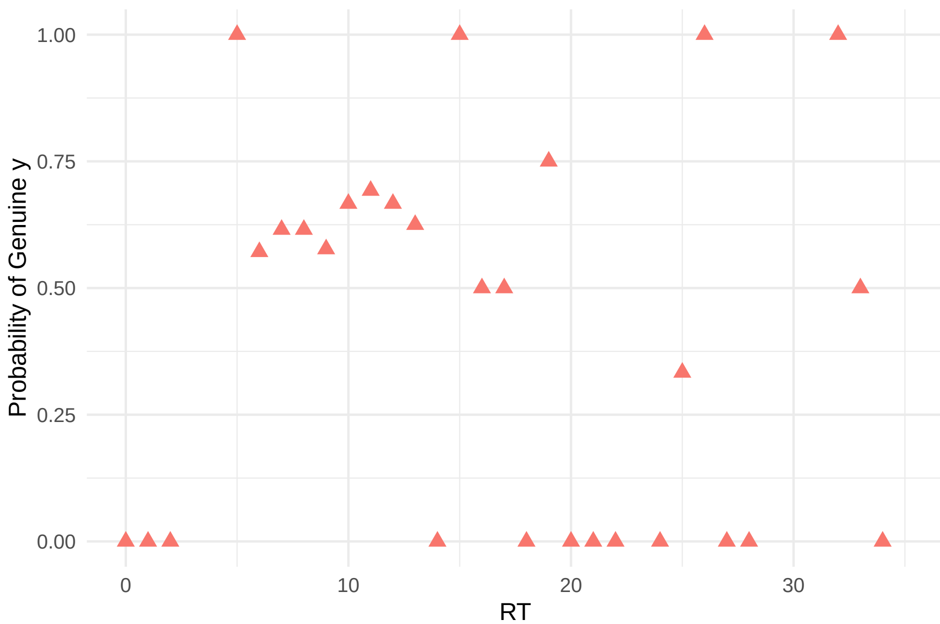
Table 2.4: Fixed Effects Regression Results: standard error in the parenthesis.

We also conducted robustness checks to further test the relationship between RT and the veracity of y reports in the SU condition. Among the 17 subjects in the SU condition who report both genuine and deceptive y, 14 display a positive relationship between RT and deception (with 3 significant at the 10% level), while 3 show a negative relationship (none significant). Additionally, we find that deceptive y reports are significantly more likely to have an RT above the sender's median RT (61%) compared to genuine "y" reports (43%), with a 20% difference significant at the 5% level. This result remains robust when controlling for the round and using clustered standard errors at the individual level.

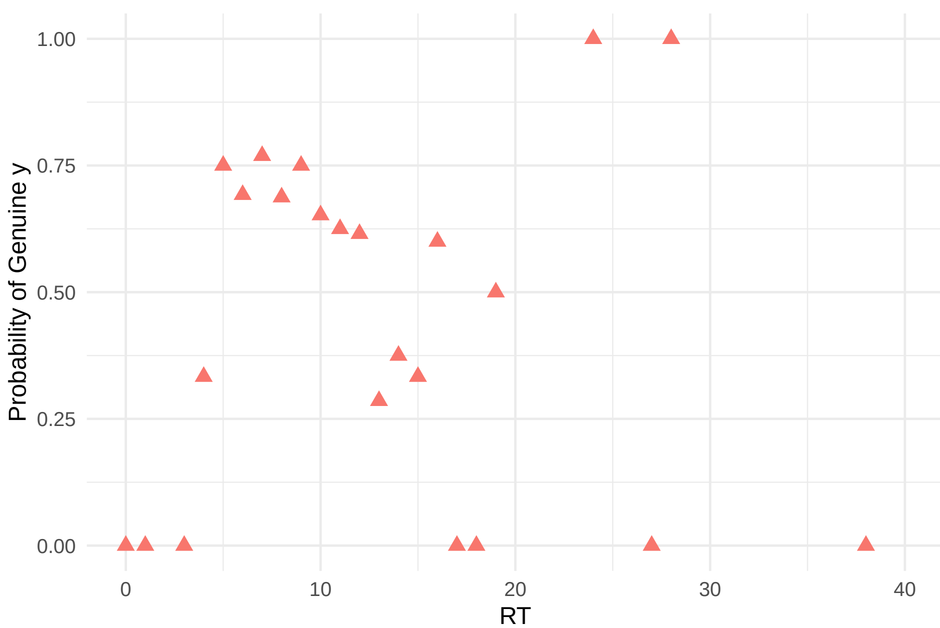
However, the fact that RT for deceptive y messages is longer than for genuine y messages does not guarantee that RT can be used as an effective cue for lie detection. RT can only serve this purpose if the probability of a longer y message being deceptive is sufficiently high. Figure 2.4a illustrates the genuine rate of y reports for each RT. To enhance the readability of the figure, we exclude two rare and extreme RT values (RT=47, and RT=90) from the analysis. A probability of 1 indicates that all y reports with that RT are genuine, whereas a probability of 0 indicates that all y reports with that RT are deceptive or uninformed. Receivers should follow the report if the honesty rate is above 50%; otherwise, they should disregard it.

The figure indicates that when the RT is within 3 seconds, no genuine y reports are observed. This finding aligns with the understanding that immediate responses are typically uninformed and self-interested. For non-immediate responses, if the RT is shorter than 20 seconds, it is more likely that the y report is genuine. Conversely, if the RT exceeds 20 seconds, it is more likely that the y report is deceptive or uninformed. These results demonstrate that RT can serve as an effective cue for detecting lies. Receivers should heed messages with non-immediate short RTs and disregard those with immediate or long RTs.

The empirical evidence contradicts the linear relationship posited in *Hypothesis 1*. Rather than a faster y report being less credible, our findings show that a y report is more likely to be genuine if the RT is non-immediate and short. Conversely, a longer RT makes the y report less credible. This result underscores the importance of understanding what RT reveals. Instead of indicating a straightforward truth-telling process, RT primarily reflects the deliberative process regarding whether to engage in deception. The finding



(a) SU Conditon. We exclude RT=47 and RT=90.



(b) SA Condition. We exclude RT=44 and RT=51.

Figure 2.4: Probability of Genuine y for Each RT

that longer RTs for informed messages are more likely to be associated with deception suggests that lying is not a hasty act and often involves considerable contemplation.

One potential explanation for the differences in RT among various messages is the variation in the number of button clicks. Our analysis revealed that for the informed reports, over 90% of participants clicked all six buttons. There were no significant differences between genuine and deceptive reports, nor between the *yes* and *no* states. Detailed statistics on the number of buttons clicked are provided in Appendix B.2. To ensure the robustness of our findings, we incorporate the number of button clicks into our analysis. The main result, indicating a relationship between RT and the type of messages, remain consistent. Here comes our second key finding.

Result 2: Immediate y reports are uninformed. For non-immediate reports, the longer the RT for a y report, the higher the probability of deceit.

Should *Proposition 1* and *Hypothesis 1* fail to hold, it naturally results in the failure of *Hypothesis 2* regarding how senders should manipulate RT in the SA condition. Given that longer RT implies a deceptive report, senders would not prolong their RT. Ideally, the sender would strategically truncate the internal struggle when lying for y, making their RT indistinguishable from genuine reports. The behavioral patterns of senders align conspicuously with this conjecture. According to Table 2.3, the average RT in the SA condition is 11.70 seconds, which is shorter than the average RT of 12.63 seconds in the SU condition. Column 5 demonstrates that almost all types of reports have shorter RTs under the SA condition compared to the SU condition. This evidence supports the notion that senders deliberately curtail hesitation to manipulate their RT.

We then examine whether senders successfully render the RT less informative, particularly for y reports. Figure 2.2b shows that the distribution of RT is more condensed in the SA condition compared to the SU condition, and there is no range of RT with more blue bars than red bars. This result implies that senders manipulate the RT of deceptive y reports to be less distinguishable from genuine y reports. Figure 2.3b supports this conclusion at the individual level, illustrating that the RTs for genuine and deceptive y reports align closely along the 45-degree line. Additionally, Column 3 in Table 2.4 presents the results of a fixed-effects regression in the SA condition, indicating that deceptive y reports are not statistically distinguishable from genuine y reports. Figure 2.4b further suggests that RT is not an effective cue for lie detection in the SA condition, with no clear criterion for when receivers should trust or disregard a message. These pieces of evidence show that senders successfully shortened deceptive hesitation and manipulated RT to be uninformative in the SA condition. Thus, we present our third result.

Result 3: In the SA condition, senders manipulated RT to carry less information than in the SU condition. This was achieved through the compression of all RTs.

Given that senders manipulate RT in the SA condition, we then study whether senders make different decisions in the two conditions, i.e., whether senders in the SA condition tell more truth as stated in *Hypothesis 3*. Table 2.3 shows that the uninformed rates among all reports and the honest rates of informed reports under the two conditions remain roughly the same. This suggests that awareness of RT does not change senders' decisions. We furhter test cumulative distribution function of the number of genuine reports among senders. Notably, the p-value derived from the Kolmogorov-Smirnov test exceeds 0.75 for both one-sided and two-sided tests. This outcome strongly illustrates a near absence of disparity in terms of reporting behavior. We present our fourth result.

Result 4: The awareness of RT availability does not alter senders' reporting decisions.

In summary, regarding the behavior of the senders, we observe a high rate of truth discovering. This scenario may be related to senders' reluctance to lie when it is unnecessary. We also observe a pattern between RT and types of reports in the SU condition. Immediate reports pertain to the uninformed selfish report, which receivers should disregard. Non-immediate short reports are associated with genuine reports, which receivers should heed, while long reports with extended RTs are associated with deceptive reports, which receivers should disregard. Lastly, when senders are aware of the availability of RT, they can successfully manipulate it to be less informative.

2.5.2 Receivers' Behavior

In this section, we examine receivers' behavior regarding the use of RT information. Firstly, we assess whether there's added value to receivers' guessing considering RT. We analyze receivers' accuracy rates in Stage 2 (RU), determining if their initial guesses are sufficiently accurate to render the RT provided in Stage 3 redundant. Alternatively, if their Stage 2 accuracy is suboptimal, it suggests potential for improvement through RT utilization. Secondly, we investigate whether receivers' RT usage aligns with rationality in response to *Result 2*, which highlights the tendency to disregard immediate y reports and those with prolonged RT.

Table 2.5 illustrates the receivers' accuracy rates across different reports. In the SU-RU condition, without RT information, receivers demonstrated a high accuracy rate of 87.8% when responding to n reports, but this dropped notably to 69.6% when countering y reports. Moving to the SA-RU condition, accuracy rates of 90.2% for n reports and 65.0% for y reports were observed, comparable to those in the SU-RU condition. The consistently modest accuracy rates for y reports highlight receivers' need for additional assistance in discerning the truthfulness of such reports across conditions. Leveraging RT could optimize decision-making by aiding receivers in distinguishing deceptive reports from genuine ones, thus enhancing overall welfare.

Sender's Report	SU-RU	SU-RI	SA-RU	SA-RI
n	88%	90%	90%	86%
y	70%	66%	65%	63%

Table 2.5: Receiver's Accuracy Rate at RU and RI

If receivers respond rationally to the information conveyed by RT, they should discount all immediate y reports with RT of less than 3 seconds across both RI conditions, according to *Result 2*. Additionally, in the SU-RI condition, receivers are advised to prioritize non-immediate y reports with RTs under 20 seconds, while disregarding those exceeding this threshold. Conversely, in the SA-RI condition, all non-immediate y reports are to be attended to, irrespective of RT. Furthermore, it is recommended that receivers consistently heed n reports, as senders' infrequent instances of downward deception. In this manner, receivers' accuracy in processing n reports within every condition should approach 99%. For y reports, the accuracy rate is anticipated to be 64% in the SU-RI condition and 62% in the SA-RI condition[4]. Table 2.5 shows that the accuracy rates in the RU conditions have surpassed the rational response thresholds of y reports. This phenomenon can be attributed primarily to receivers' astute recognition of the heterogeneity among senders, leveraging such variability to optimize outcomes. We leave the heterogeneity analysis to the next section, and we focus on the general reaction in this

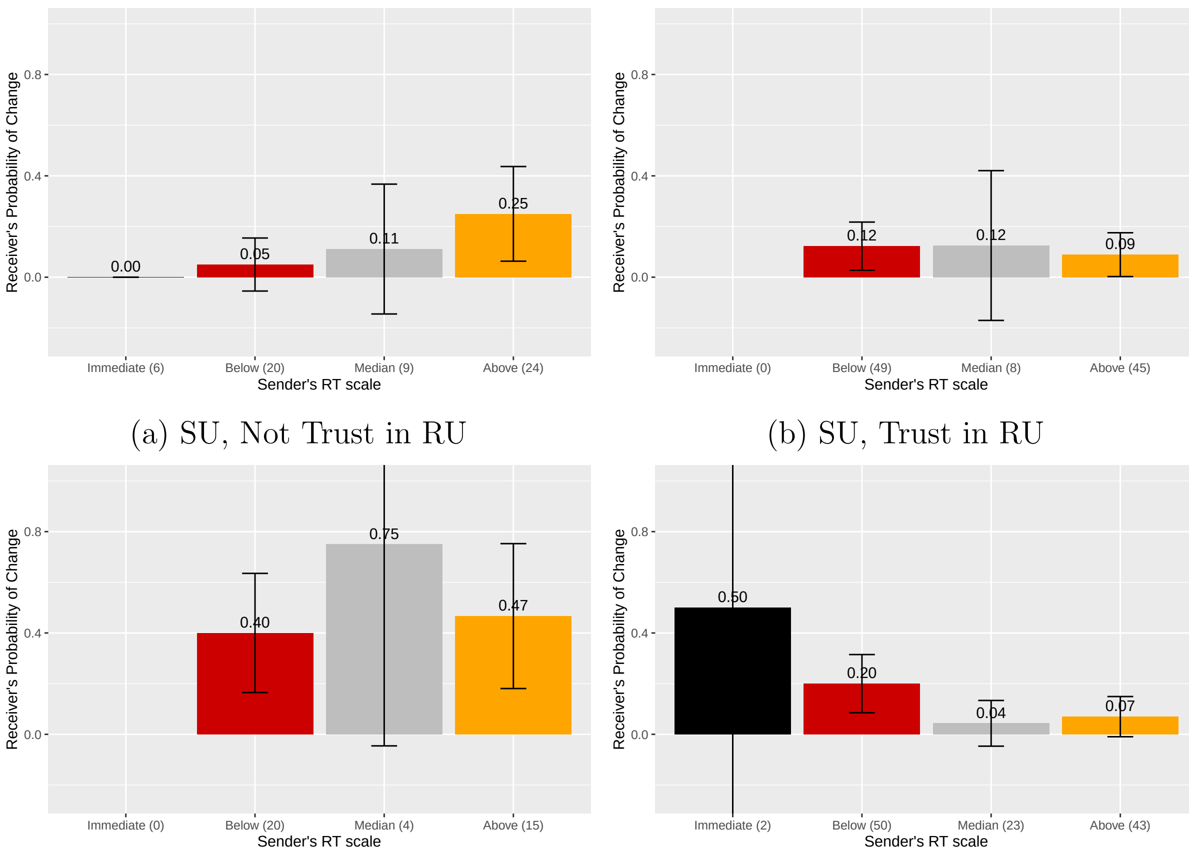
[4]The anticipated accuracy rate is calculated as the weighted average of the accuracy rate for guessing YES when the RT of the y report falls between 4 and 20 seconds in the SU condition, and responding all other reports by NO, weighted by the frequency of each type of report. We simplify by assuming a 100% correct rate for guessing immediate reports as NO, and that the accuracy rate for guessing non-immediate reports aligns with the genuine rate. However, it's essential to acknowledge that this calculation deviates slightly from the precise accuracy rate due to inherent uncertainties in guessing. For instance, there is a 34% chance that the immediate y report is associated with the state yes, thus rendering guessing NO inaccurate. Furthermore, the possibility exists that the non-immediate y report is the uninformed report, where guessing NO and YES can be either right or wrong.

section.

The accuracy rates observed between RU and RI conditions point to receivers' utilization of RT as a pivotal determinant in their decision-making framework. Subsequently, our inquiry focuses on the intricate mechanisms through which receivers harness RT for decision-making. To ensure methodological robustness, we classify responses into four discrete strata: immediate reports (RT=0), reports with RT falling below the individual non-immediate median, reports with RT at the median, and reports with the RT surpassing the median threshold.

Figure 2.5 illustrates receivers' change rates for y reports. Panel (a) and (b) show the change rates for different initial guess in Stage 2 (RU) in the SU condition. Contrary to the rational shift, receivers shifted from NT to T for reports with Above RT significantly more than ones with Below RT (p=0.061), while there is no significant difference in change rates between Above and Below RT when they changed from T to NT. This suggests a tendency for individuals to place more trust in reports with longer RT compared to those with shorter RT. Panel (c) and (d) portray the change rates in the SA condition. While there is no significant difference for reports with Below and Above RT when receivers changed from NT to T, they shifted from T to NT for reports with Below RT significantly more than ones with Above RT. Similarly as the behavior in the SU condition, receivers trust the long RT in the SA condition as well.

Given that receivers changed their guesses for n reports after knowing RT, as opposed to always trusting them, we'd love to explore the mechanisms receivers use for n reports. Given that receivers trust n reports at around 90% in both conditions, we have less than 20 data points for each categories when receivers initially chose NT. Considering that there are not enough data to study the change from NT to N, we focus on the change from T to NT for n reports. Figure 2.6 illustrates receiver's change rates for n reports when they initially chose T in Stage 2 (RU). Panel (a) shows that receivers are significantly



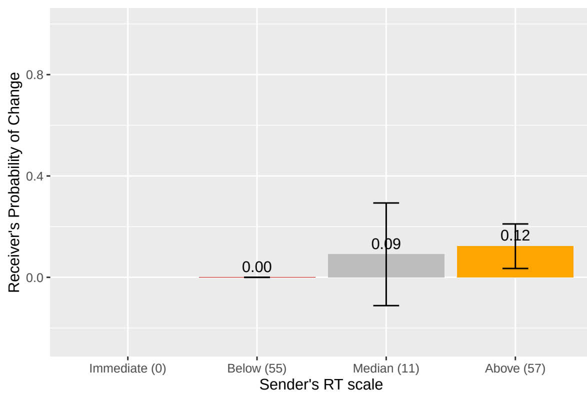
(a) SU, Not Trust in RU

(b) SU, Trust in RU

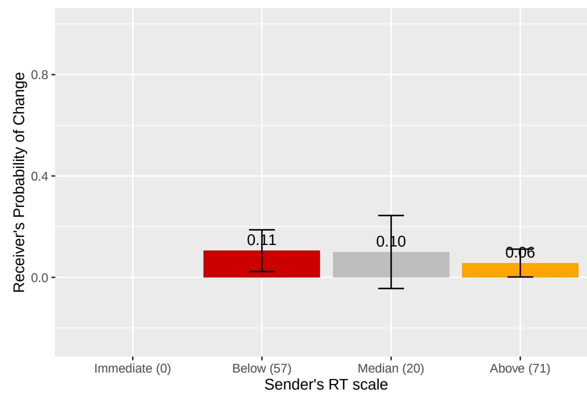
(c) SA, Not Trust in RU

(d) SA, Trust in RU

Figure 2.5: Receiver's Change Probability given Sender's RT for y Reports. The difference between the Above and Below is significant in (a) and (d) at 10% level; the number on the bar denotes the mean of honesty rate, the two smaller bars represents the 95% confidence interval.



(a) SU, Trust in RU



(b) SA, Trust in RU

Figure 2.6: Receiver's Change Probability given Sender's RT for n Reports. The difference between Above and Below is significant in (a) at 1% level; the number on the bar denotes the mean of honesty rate, the two smaller bars represents the 95% confidence interval.

more likely to turn to distrust reports with Above RT than with Below RT (p=0.007) in the SU condition. Panel (b) portrays that there is no such difference in the SA condition. In general, different from attitudes for y reports, receivers distrust n reports in the SU condition.

Overall, contrary to the discernible pattern between the RT and credibility, receivers tend to trust long non-immediate y reports and short non-immediate n reports in the SU condition. They rely on RT less in the SA condition. For the robustness check, we ran the OLS regression by controlling the individual and round fixed effect, the results remain the same (see results in Appendix C.2). We now draw our conclusion for receiver's behavior.

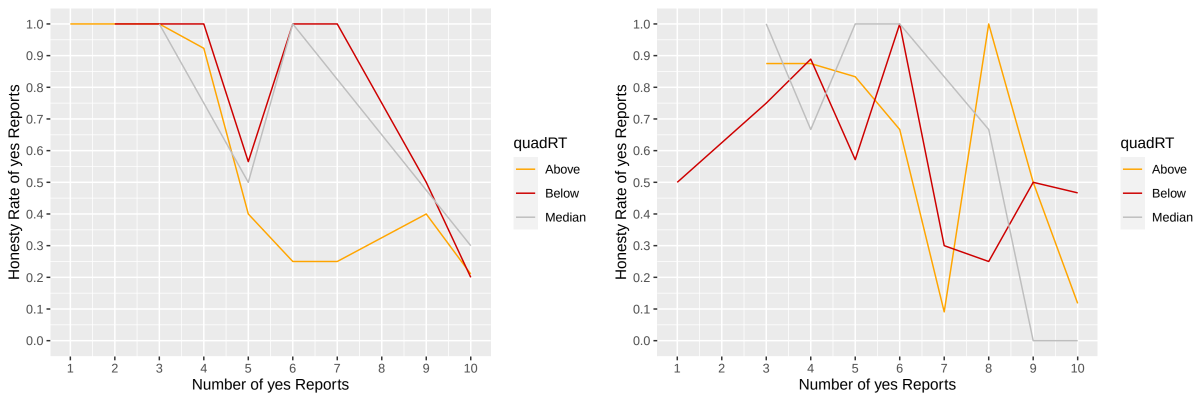
Result 5: In the SU condition, receivers trust the short y than the long y, and they distrust the long n than the short n. Receivers rely less on RT in the SA condition.

Given the insufficient use of RT in the SU condition, receivers' welfare do not increase after knowing RT in the SU condition (p=0.84 in two-tailed test). There was no significant change in receivers' welfare between the SU and the SA condition (p=0.35). As a result of receivers' behavior, there ware no significant differences in senders' expected payoffs after revealing their RT in both conditions (p=0.61 in SU, p=0.52 in SA). See more details of payoff comparisons in Appendix C.1.

2.5.3 Heterogeneity

The relatively high accuracy rates among receivers suggest that they use other information beyond mere RT for guessing. Our initial focus centers on the frequency of y reports over 10 rounds. As depicted in Panel (a) in Figure 2.7, the genuity rate of y reports varies concerning different non-immediate RT intervals and the quantity of y reports in the SU condition. Notably, there is added value by RT particularly when the tally of y reports falls within the 5 to 9 range over the 10 rounds. Within this span, reports characterized by Below Median hold superior credibility compared to their counterparts with Above Median. Therefore, receivers are supposed to accord greater weight to reports with shorter RTs while discounting those with elongated RTs. Subsequently, Panel (b) of the figure illustrates that RT loses its informative capacity within the SA condition across various senders, thus highlighting its limited utility in such contexts.

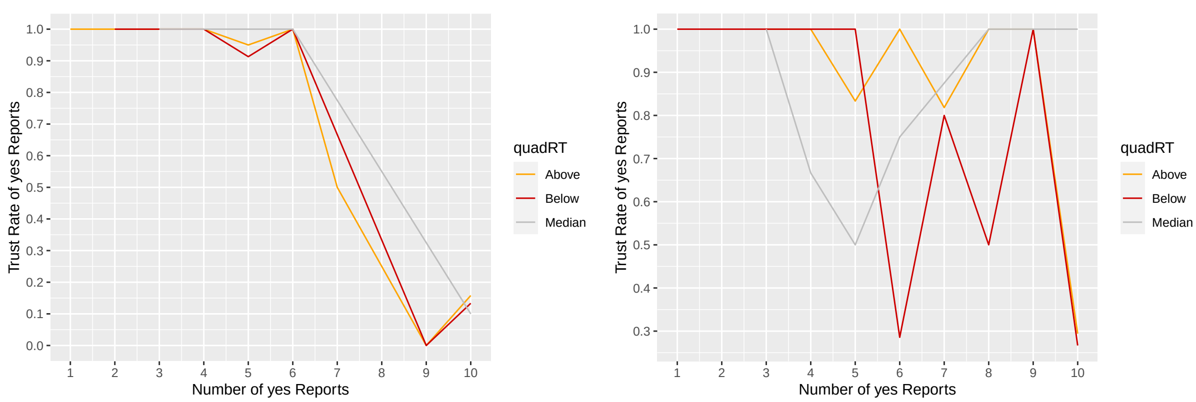
Given the multifaceted nature of sender heterogeneity, our investigation proceeds to examine how receivers navigate and respond to such diversity. Figure 2.8 offers insights into receivers' trust dynamics within the RU condition, operating under the absence of sender RT information. Panel (a) illuminates that receivers, while not discerning among RTs, exhibit a nuanced response to the quantity of y reports. Notably, they display a propensity to place greater trust in y reports when the count remains below 6, while exhibiting diminished trust as the number of y reports surpasses this threshold, particularly in conjunction with prolonged RTs. Transitioning to Panel (b), a similar



(a) SU condition (b) SA condition

Figure 2.7: Heterogeneous Honesty Rate of Non-immediate y Reports Given Different Number of y Reports.

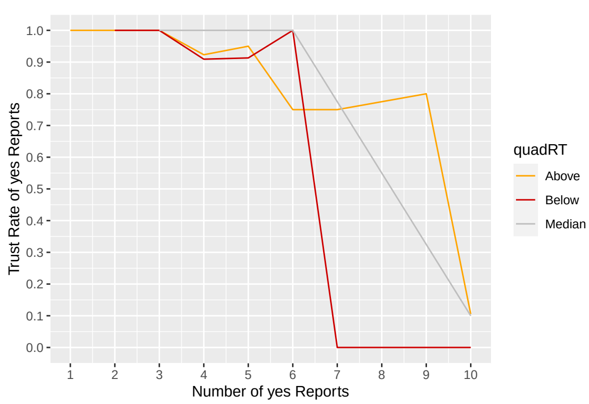
trend emerges within the SA condition, albeit with heightened variability in receivers' responses to RT. This heightened variability suggests a more intricate interplay between RT and trust within the SA context. Overall, these findings underscore that receivers respond rationally to the number of y reports.



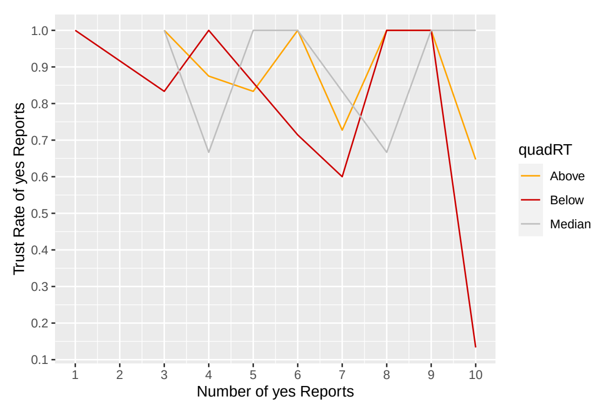
(a) SU-RU condition (b) SA-RU condition

Figure 2.8: Trust Rate of Different RT Scales for y Reports in RU conditions

Figure 2.9 sheds light on receivers' trust rates within the RI conditions subsequent to acquiring sender RT information. Panel (a) reveals a noteworthy divergence from expected rational behavior in the SU condition: receivers exhibit a higher degree of trust in y reports characterized by longer RTs compared to those with shorter RTs, when the tally of y reports falls within the range of 7 to 9. This departure from rational expectations stands in stark contrast to the behavior observed in Panel (a) in Figure 2.7. Conversely, in the SA condition, receivers display a propensity to rarely factor RT into their decision-making process. These findings highlight a notable irrationality among receivers in the SU condition, wherein there is a tendency to over-trust reports with extended RTs. One plausible explanation for this phenomenon could be rooted in receivers' perceptions regarding the informational value of RT. It's conceivable that receivers associate prolonged RTs with a higher likelihood of senders uncovering the truth, thus attributing greater accuracy to such reports. However, the underlying rationale driving this bias warrants deeper investigation.



(a) SU-RI condition



(b) SA-RI condition

Figure 2.9: Trust Rate of Different RT Scales for y Reports in RI conditions

We then delves into an exploration of the potential impact of gender and creativity traits on subjects' behavior as stated in *Hypothesis 5*. These attributes are assessed through self-report methodologies in the post-experiment survey. Specifically, our attention centers on the veracity of disclosures, the inclination towards trust, and the resultant payoffs. The descriptive statistics are synthesized in Table 2.6, detailing the results concerning gender and creativity across two conditions. A meticulous examination of the data reveals a lack of significant disparities in gender, but subjects under SU condition have higher creativity score than those under the SA condition (p = 0.03). Given the small observations under each treatment, we pool the data together for analysis.

	SU	SA
Gender	M 11, F 18, O 1	M 11, F 20, O 1
Creativity: mean(s.e.)	3.83(0.59)	2.09(0.51)

Table 2.6: Gender and Creativity Statistics. M for male, F for female, O for other. The creativity score is measures by the Gough personality scale, ranging from -12 to 18. A higher score indicates greater creativity.

Commencing with an analysis of gender-based heterogeneity in behavior, we observe notable distinctions. Specifically, in terms of the honesty rate, male subjects exhibited a truth-telling rate of 64.5%, whereas their female counterparts demonstrated a heightened rate of 82.1%. This significant discrepancy (p = 0.04) underscores a gender-related variance in truthfulness, elucidating the propensity of female subjects to lean towards honesty within the experimental framework.

Turning to the assessment of trust rates, our examination reveals small and insignificant gender-related patterns. For male subjects, the trust rate registered at 74.1% in the absence of RT (RU) and remained fairly consistent at 73.6% when RT is informed (RI). Female subjects manifested trust rates of 81.1% without RT (RU) and 80.8% with RT (RI). The overall disparity magnitude is around 7%, which is about one choice difference. However, the observed inter-gender differences within each condition do not attain statistical significance ($p > 0.40$ for both RU and RI conditions). In light of these findings, it becomes evident that females leaned slightly more, though not significant, towards an overarching inclination for trust.

Examining expected payoffs from the sender's perspective, male subjects achieved an anticipated payoff of $5.50 under both the RU and RI condition, whereas female subjects secured an expected payoff of $5.70 under both the RU and RI conditions. Female senders' slightly higher expected payoff, though not significant ($p > 0.3$ for each condition) suggests that females benefited from telling more truth. Shifting to the expected payoffs in the role of receivers, male subjects realized \$7.30 under the RU condition and \$7.20 under the RI condition, whereas their female counterparts received \$7.00 and \$6.90, separately. Remarkably, the disparity by gender proves significant ($p = 0.10$) under the RU condition, while it loses significance ($p > 0.2$) under the RI condition. This finding suggests that males gained advantage from suspicion, particularly when they are devoid of the opportunity to use RT. We now wrap up the overall gender heterogeneity.

Result 6: Females tend to tell more truth and trust others more. Males gain advantage from suspicion as a receiver, when they are devoid of the opportunity to exploit RT.

We subsequently investigate the impact of creativity on various outcomes. Creativity is quantified using the Gough Scale [Gough, 1979]. According to the protocol, 1 point is given each time one of the 18 positive items is checked, and 1 point is subtracted each time one of the 12 negative items is checked. The theoretical range of scores is therefore from –12 to +18. A higher score indicates greater creativity. To empirically examine this relationship, we apply a regression framework 2.2:

$$Y_i = a + b \times creativity_i + \epsilon_i \tag{2.2}$$

The dependent variable Y_i comprises metrics including sender i's honesty rate, trust rate under both RU and RI conditions, and payoffs across four distinct role categories ((Sender, RU), (Sender RI), (Receiver, RU), (Receiver, RI)).

Table 2.7 presents the outcomes of the regression analysis. The presence of statistically insignificant coefficients across all outcomes implies that creativity, particularly when measured through self-report assessments, does not demonstrate a notable impact on the rates of truth-telling, trust, or the resulting payoffs within the experimental framework.

Result 7: There is no observed influence of self-reported creativity on the rates of truth-telling, trust, or the subsequent payoffs.

2.6 A Calibrated Utility Function

We calibrate the senders' utility function to better understand their behavior. We formulate the senders' utility function by incorporating their reluctance to lie:

$$u_S(f(d), a, c(\omega, m; \theta)) = (n_I - c_I)\mathbb{1}_{d=1} + h\mathbb{1}_{a=\omega_H} - l\mathbb{1}_{a=\omega_L} - c_\theta \mathbb{1}_{m \neq \omega}$$

.

The monetary payoff and the cost of lying remain the same as before. But now, beyond the discovering cost c_θ, there is a fixed benefit of knowing whether it is necessary to lie n_I by getting informed of the truth. The fact that the vast majority of the senders uncover the truth suggests that the net value from discovering the truth is positive, i.e., $n_I - c_I \geq 0$ for the majority of people. In that case, there are babbling equilibrium, truth telling, and partial lying equilibrium.

This utility function could also explain the fast and slow patterns based on the drift diffusion model. The drift diffusion model claims that the larger the difference of the choices, the faster the decision. From now on, we assume that the sender chooses to discover the truth, and the receiver trusts any message from the sender with a positive

	Dependent variable:						
	honesty	trust RU	trust RI	SProfit RU	SProfit RI	RProfit RU	RProfit RI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
creativity	0.080	0.017	0.014	-0.027	-0.047	-0.027	-0.010
	(0.118)	(0.130)	(0.120)	(0.037)	(0.034)	(0.032)	(0.033)
Constant	7.361***	7.868***	7.846***	5.699***	5.771***	7.196***	7.067***
	(0.507)	(0.561)	(0.515)	(0.159)	(0.147)	(0.139)	(0.142)
Observations	62	62	62	62	62	62	62
R^2	0.008	0.0003	0.0002	0.009	0.031	0.012	0.001
Adjusted R^2	-0.009	-0.016	-0.016	-0.008	0.015	-0.005	-0.015
Note:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 2.7: Creativity Results

probability $p > 0$, which explains 83% of the data[5].

Let's first compare the genuine y and the genuine n reports. Given a *yes* state, the difference between reporting y and n is $D_1 = pz + c_\theta > 0$. For any realization of c_θ, the sender would always report y. Given a *no* state, the difference between reporting y and n is $D_2 = pz - c_\theta$. The sender would report y if $pz > c_\theta$, and otherwise n. No matter what the sender decides to do, the absolute value of D_2 is smaller than D_1. Therefore, both the deceptive y and the genuine n are slower than the genuine y.

2.7 Conclusion

We study what private information RT carries and its strategic use in a sender-receiver game with conflict of interest and with a truth discovery process. We found that the vast majority of people discovers the truth before deciding to lie or to tell the truth. It implies that RT does not reveal if the sender uncovers the truth. Given such strong preference of getting informed of the truth, we have found that for the informed messages with a conflict of interest, the longer the message, the higher probability of deceit. However, receivers did not use it effectively. When receives had better trust more relative short messages, they did trust more long messages. We also found that senders were able to successfully manipulate their RT to their best interests, and availability of RT did not change their honesty rate.

These findings suggest that long RT is not an effective cue in detecting lies, which is also easily to be manipulated. Trusting toward the long RT is not an rational decision for receivers. Receivers might be better off if they just ignore the RT and trusting the message more when it's against the sender's interest.

One limitation of our research is that the scale of lying benefit and cost of discovering

[5] Senders uncovered the truth 92% of the time, and receivers trusted each message with positive probability 90% of the time.

might be too small to make an big influence in senders' decision. In real life, the benefit of recommending an expensive car might be worth a month's salary and cost of meticulously examining each corner of a house might be equivalent to several hours' pay. Future research could study senders' lying and discovering behavior with a high lying benefit and a high discovery cost. There are also environment when lying is supposed to be costly, and future research could investigate the strategic use of RT in such environment.

Chapter 3

Dynamic Binary Elicitation Method

Joint work with Jing Zhou

3.1 Introduction

Information on individual beliefs is central for researchers to better understand economic behavior [Manski, 2004]. Without data on what people think and expect, it is challenging to differentiate between alternative choice models, understand the boundaries of rationality, or examine new equilibrium concepts. However, eliciting individual beliefs poses its own set of challenges. Existing methods primarily rely on individuals selecting a number from 0% to 100% to represent their probabilistic beliefs [Charness et al., 2021], which introduces numerous issues. For instance, individuals may possess imprecise rather than precise probabilistic beliefs about a particular event [Giustinelli et al., 2022]. They might have a general notion but struggle to provide the best point estimate. When asked to state a point belief, cognitive difficulties may arise, leading to conservative responses and systematic deviations from truthful reporting [Charness et al., 2021].

In this paper, we introduce a new elicitation method called the *Dynamic Binary Method* (DBM). Unlike *Classical Methods* (CM), which directly ask respondents to select a number from 0 to 100 (%) as their probabilistic beliefs, and use proper scoring rules such as the Binary Scoring Rule (BSR) to incentivize truthful reporting, DBM differs in how beliefs are stated and whether they must be a single value. Inspired by the bisection process – the iterated partition of a choice set into two equally large subsets, with perceptions elicited through a series of binary choices [Baillon, 2008] – DBM allows respondents to exit at any step and state interval beliefs if they prefer. This method is designed to achieve two primary goals: (1) alleviate the challenge of forming a precise point estimate of beliefs or perceptions, and (2) quantify the self-perceived precision of those beliefs.

To elaborate further, starting with the full belief space, in each step s, DBM divides the belief space $[I_l^s, I_u^s]$ into two equally sized intervals: $[I_l^s, \frac{I_l^s+I_u^s}{2}]$ and $(\frac{I_l^s+I_u^s}{2}, I_u^s]$, where I_l^s and I_u^s denote the lower and upper bounds of the presented interval, respectively. The decision maker (DM) must then select either $[I_l^s, \frac{I_l^s+I_u^s}{2}]$ or $(\frac{I_l^s+I_u^s}{2}, I_u^s]$, or they can opt to exit with the interval $[I_l^s, I_u^s]$. If the DM chooses to exit, the computer randomly selects a number a_R from the stated belief range $a = [a_l, a_u]$, following a uniform distribution. The selected number a_R is then applied in a proper scoring rule, such as the BSR, to determine the DM's payoff.

For an expected utility maximizer, choosing the mean of their true belief, no matter whether their true belief is precise or distributed, is optimal in both DBM and CM. However, an expected utility maximizer who does not perfectly foresee the optimal choice but instead considers randomization over $[I_l^s, \frac{I_l^s+I_u^s}{2}]$, $(\frac{I_l^s+I_u^s}{2}, I_u^s]$, or $[I_l^s, I_u^s]$, may opt to exit early with an interval whose midpoint equals the mean of their true belief. Thus, the decision to exit early indicates whether the DM is myopic or not. DBM also facilitates relative judgment by asking which range is more likely, thereby sidestepping the challenge of finding the best point estimate. If it is main driving force behind biases in perception, for example, the compressed relationship between respondents' probabilistic estimates and "true" probabilities, it would have the potential to mitigate the difficulty of forming precise point estimates.

To assess the empirical validity of DBM, we conduct both within-subject and between-subject experiments using a diverse range of perception tasks from previous literature. Specifically, for the between-subject design, we utilize four task categories from controlled laboratory experiments: simple prior tasks [Danz et al., 2022], compound prior tasks [Liang, 2022], belief updating tasks [Danz et al., 2022], and estimation tasks [Dewan and Neligh, 2020, Falk and Zimmermann, 2018] with artificial settings such as balls and urns, counting peas in a bowl, or dots in a graph. Additionally, we incorporate four task categories from field or lab-in-the-field experiments: perception on economic or financial variables [Enke and Graeber, 2023], the labor market [Wiswall and Zafar, 2015b], and education [Wiswall and Zafar, 2015a], all of which have real-life settings.

To address the challenge of not knowing participants' true beliefs, we carefully design the questions to ensure that each task has an objective truth. Furthermore, we intentionally select parameters for each question to ensure that the objective truths span the entire belief space, including centered, extreme, and intermediate values. For the within-subject design, we allow each participant to complete a set of perception tasks using both DBM and CM in a randomly determined order. This approach aims to assess the extent to which the elicited beliefs in tasks using DBM can predict stated point beliefs in tasks using CM at the subject level.

First of all, we find that DBM does not perform significantly differently from CM at the aggregate level, regardless of whether the perception questions use artificial/laboratory settings or real-life settings. This finding is robust across different measures, including the absolute deviation of the midpoint of elicited beliefs from the objective truth or the expected absolute deviation of elicited beliefs from the objective truth. This suggests that the challenge of forming a point estimate of beliefs/perceptions may not be the primary driver of biased perception elicited using CM. But DBM outperforms CM when the task has extreme values as the objective truth. This implies that some perception biases, such as central tendency, could result from the narrowed consideration set that respondents use to choose beliefs or perceptions from.

Furthermore, we find, from both between-subject and within-subject perspectives, that the length of stated beliefs in tasks using DBM is negatively correlated with their accuracy and can predict how well respondents perform in tasks using CM at the aggregate level: the longer the interval, the less accurate the stated belief in DBM and the less accurate the stated belief in CM. Moreover, within-subject results highlight participants' sophistication regarding the precision of their beliefs/perceptions: participants who stated point beliefs in DBM in more tasks demonstrate less deviation from the objective truth in their stated beliefs in CM. This pattern is particularly significant among participants who completed tasks with DBM first and subsequently used CM.[1]

Note that this relationship is not strictly monotonic: stated beliefs reaching the point are not the most accurate and do not predict the most accurate beliefs stated in CM. Participants who always choose until reaching the point in all tasks using DBM are not the most accurate in tasks using CM. Moreover, our findings reject the hypothesis that participants have precise beliefs/perceptions but do not bother to choose until reaching the points for reasons such as complexity. If this were the case, we would expect no correlation between the length of their stated beliefs in DBM and the absolute deviation of their stated point beliefs from the objective truth in CM. This finding suggests that participants possess some level of awareness regarding how accurate their beliefs/perceptions

[1] We interpret this difference as a fatigue effect as in our Experiment 1, subjects are underpaid given the time they took to finish the experiment and the standard payment suggested by Prolific.

would be when using DBM.

Lastly, we compare three methods of using the stated beliefs elicited with DBM to predict point beliefs elicited with CM. We find that predictions using a weighted average between subjective truth (the midpoint of stated beliefs in DBM) and the cognitive default (e.g., midpoint of the slider bar), with the relative weight on the default determined by the length of stated beliefs in DBM, are closest to the average stated beliefs in CM. This approach outperforms both using the midpoint of stated beliefs in DBM alone and using objective truth instead of subjective truth in the weighted average method. Our findings underscore the significance of incorporating the precision of stated beliefs and perceived truth to enhance predictions of economic behavior.

Relations to the existing literature. This paper makes several contributions to the existing literature. Firstly, our study aligns closely with previous research on perception/evaluation imprecision and the notion of cognitive uncertainty introduced by Enke and Graeber [2023]. Most studies in this domain focus on capturing preference incompleteness, cognitive noise, or cognitive uncertainty using non-incentivized techniques. For instance, Enke and Graeber [2023] measure "cognitive uncertainty" by having participants first choose from a slider bar to state their beliefs/perceptions and then report a probabilistic value indicating the extent to which they are "certain" about their previous choice is the best on a second screen without incentivizing truth-telling. Similar technique is used in Giustinelli et al. [2022], Nielsen and Rigotti [2023] for the identification of belief imprecision by asking participants to report probability intervals after the question using a precise percent-chance format, with the question about belief range being unincentivized. Recently, Agranov and Ortoleva [2020] proposes an incentivized method to measure the extent to which people choose to randomize between two risky options,focusing on eliciting the ranges of preference for randomization in the domain of choice under risk.

Our study contributes to the literature by proposing a new incentivized method for eliciting participants' imprecise beliefs in the domains of perception and inference.

Secondly, our study is situated within the growing empirical literature on preferences from randomization. Existing studies have documented randomization in various contexts, including objective lotteries [Agranov and Ortoleva, 2017, Dwenger et al., 2018, Feldman and Rehbeck, 2022], ambiguity preferences [Cettolin and Riedl, 2019], time preferences [Agranov and Ortoleva, 2017], social preferences [Agranov and Ortoleva, 2017, Miao and Zhong, 2018], and even choices involving dominated options [Agranov et al., 2023, Rubinstein, 2002]. The survey paper by Agranov and Ortoleva [2022] demonstrates high rates of preferences for randomization across these domains and shows their persistence even after explicit training.

Similar to these studies, we capture the prevalence of randomization using incentivized measures. Moreover, we extend this line of inquiry into the domain of belief formation and inference and document the prevalence of randomization over beliefs, thereby complementing existing literature in this area.

The rest of this paper is organized as follows. Section 3.2 delves into the theoretical benchmark of DBM and CM with BSR. Section 3.3 outlines the experimental design. Section 3.4 presents the results, and Section 3.5 concludes.

3.2 Theoretical Framework

Consider a decision maker (DM) with a probabilistic belief over a verifiable binary outcome $s \in \{A,\ B\}$, assuming they possess a true belief $p = Pr\{s = A\}$. Binarized scoring rule (BSR) uses two monetary prizes M_h and M_l for payment (where $M_h > M_l \geq 0$), and two i.i.d. draws $X_1, X_2 \sim U[0, 1]$ to determine the outcome [Hossain and Okui, 2013, Wilson and Vespa, 2018]. Specifically, if $s = A$ is true, the DM gets the prize M_h so long as their stated belief a is greater than at least one of the two uniform draws X_1 and X_2. If $s = B$ is true, the DM gets the prize M_h so long as their stated belief a is less than at least one of the two uniform draws X_1 and X_2. Otherwise, the DM gets the prize M_l. Given the true belief p, the probability of winning the better prize M_h with the stated belief a is

$$\pi(p, a) = p * (1 - (1 - a)^2) + (1 - p) * (1 - a^2) \tag{3.1}$$

Thus, BSR generates a reduced lottery $\mathcal{L}(a|p)$ that getting the prize M_h with the probability $\pi(p, a)$ and M_l with the probability $(1 - \pi(p, a))$, such as:

$$\mathcal{L}(a|p) = \pi(p, a) \circ M_h \oplus (1 - \pi(p, a)) \circ M_l \tag{3.2}$$

Without loss of generality, assume $M_l = 0$. Given the true belief p, finding the optimal stated belief $a \in [0, 1]$ that maximizes the expected utility in the BSR is equivalent to maximizing the likelihood of receiving the prize M_h.

Classical Methods (CM) refer to implementation methods that elicit the DM's stated point belief a by directly asking the DM to report any value within the full choice space, such as any real number between 0 and 1. As this approach is widely used in existing literature, we refer to them as Classical Methods (CM). The stated belief a is then used in Equation (3.1) to determine the lottery for their rewards, i.e., Equation (3.2), and the outcome is realized accordingly.

Dynamic Binary Method (DBM) is based on the bisection method. The choice interval is repeatedly partitioned into two equally lengthy sub-intervals, for which the DM's beliefs are elicited through a series of binary choices. Starting with the full choice space, for example, $[0,1]$, at each step, the method divides the choice interval $[I_l, I_u]$, where I_l and I_u denote the lower and upper bounds separately, in two halves at the midpoint $\frac{I_l+I_u}{2}$: $[I_l, \frac{I_l+I_u}{2}]$ and $(\frac{I_l+I_u}{2}, I_u]$. Unlike the standard bisection method, which requires the DM to continue until a specific point is reached, DBM allows the DM to exit at any step and choose the current interval $[I_l, I_u]$ as their belief. Upon exiting the process, a_R is randomly selected by the computer within the last range $[I_l, I_u]$, following a uniform distribution. This a_R is then used as the stated belief in the BSR to determine the lottery for their rewards, i.e., Equation (3.2), and the outcome is realized accordingly. Therefore, the DM can either choose until the point where $a = I_l = I_u$, or select an interval $[I_l, I_u]$ as their stated belief, with a_R uniformly distributed within this interval.

3.2.1 Incentive Compatibility with CM for EU Maximizer

When the true belief p is precise, i.e., a singleton, and the CM is employed to elicit the stated belief, with BSR, the best response is to choose the point where $a^*(p) = p$ because $\mathcal{L}(a^*(p)|p)$ stochastically dominates any other available lottery $\mathcal{L}(a|p)$ [Wilson and Vespa, 2018]. Conversely, when the true belief p follows a non-degenerate distribution $f(p)$ with the mean $\mu_p = E(p)$ and variance $\sigma_p^2 = Var(p) > 0$, and the CM is used to elicit the stated belief, the objective becomes maximizing the expected likelihood of receiving the prize M_h:

$$\max_a \; E_p[p * (1 - (1-a)^2) + (1-p) * (1 - a^2)] \tag{3.3}$$

The distribution over p reflects the idea that the perception of $Pr\{s = A\}$ can be noisy, uncertain, or imprecise [Enke and Graeber, 2023, Frydman and Jin, 2022, Giustinelli et al., 2022]. The best response in this situation is to select the point $a^*(p)$ where $a^*(p) = E(p)$.

Proposition 1. *Given the true belief p, regardless of whether the true belief is a singleton or a distribution, when the CM is used to elicit belief as a singleton and BSR is used to determine payoff, an expected utility maximizer will choose the point $a^*(p)$ where $a^*(p) = E(p)$.*

3.2.2 Incentive Compatibility with DBM for EU Maximizer

Since the DBM allows the DM to either continue until reaching a single point or exit early with a random variable uniformly distributed over the last range they chose, i.e., $a \sim Uniform[a_l, a_u]$, the optimization problem becomes:

$$\max_a \; E_p\{p * E_a[(1-(1-a)^2)|p] + (1-p) * E_a[(1-a^2)|p]\} \tag{3.4}$$

where a can either be an interval $a \sim Uniform[a_l, a_u](a_l \neq a_u)$ or an point $a = a_l = a_u$. The maximization problem is equivalent to

$$\max_a \{-Var(a) - [E(a) - E(p)]^2 + E(1-p) + [E(p)]^2\} \tag{3.5}$$

where $Var(a)$ and $E(a)$ denote the variance and the mean of stated belief a, respectively. To maximize the expected utility, it is optimal to choose until the point a^* where $a^* = E(p)$ and $Var(a) = 0$.

In sum, given the true belief p, to maximize expected utility, it is optimal to continue until reaching the point $a^*(p) = E(p)$. This holds true regardless of whether the true belief follows a non-degenerate distribution or whether the DM is forced to choose a single point as their belief.

Proposition 2. *Given the true belief p, when the DBM is used to elicit belief without forcing the DM to choose a single point as their belief and BSR is used to determine payoff, an expected utility maximizer will choose until reaching the point $a^*(p) = E(p)$.*

3.2.3 Incentive Compatibility with DBM for Myopic EU Maximizer

If the DM is myopic, instead of perfectly foreseeing that the optimal choice is to choose until the point a^* where $a^* = E(p)$ and $Var(a) = 0$ in the DBM, they may compare among the three options, followed by an immediate exit in each step: choosing $Uniform(I_l, \frac{I_l+I_u}{2})$, choosing $Uniform(\frac{I_l+I_u}{2}, I_u)$, or choosing $Uniform(I_l, I_u)$. Whenever $E(p) < \frac{I_u+I_l}{2}$, choosing $Uniform[I_l, \frac{I_l+I_u}{2}]$ yields a higher likelihood of receiving M_h than choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$ or exiting with $Uniform[I_l, I_u]$. Similarly, whenever $E(p) > \frac{I_u+I_l}{2}$, choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$ yields a higher likelihood of receiving M_h than the other two options. Whenever $E(p) = \frac{I_u+I_l}{2}$, all three options yield the same likelihood of receiving M_h. Thus, the DM would be indifferent in choosing any of the three options.[2] Detailed proof can be found in Appendix D.2. In other words, whenever $E(p)$ is strictly within one of the two narrowed intervals, it is optimal to choose the one that contains $E(p)$. Otherwise, the myopic DM is indifferent between choosing $Uniform[I_l, \frac{I_l+I_u}{2}]$, choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$, or exiting with $Uniform[I_l, I_u]$.

Proposition 3. *If the DM is myopic – fails to perfectly foresee that choosing until $a^* = E(p)$ is optimal in the DBM, then, in each step, they will be indifferent among $[I_l, \frac{I_l+I_u}{2}]$, $(\frac{I_l+I_u}{2}, I_u]$, or exiting with $[I_l, I_u]$, whenever $E(p) = \frac{I_l+I_u}{2}$. Otherwise, it is optimal to always choose the interval which strictly contains $E(p)$.*

[2] For continuous uniform distribution, whether $E(p) = \frac{I_l+I_u}{2}$ is contained in the left interval or right interval does not matter. For discrete uniform distribution, the myopic DM will be indifferent among $[I_l, \frac{I_l+I_u}{2}]$, $[\frac{I_l+I_u}{2} + 1, I_u]$, or exiting with $[I_l, I_u]$, whenever $E(p) = \frac{I_l+I_u}{2} + \frac{1}{2}$.

In sum, in each step, DM will choose the interval contains the mean of their true belief. They exit whenever the midpoint of the interval (point) is equal to the mean of DM's true belief. Additionally, without further behavioral assumptions, early exit in the DBM indicates whether the expected utility maximizer is myopic – fails to perfectly foresee that choosing until $a^* = E(p)$ is optimal. This is orthogonal to whether their true belief is a precise singleton or an imprecise interval.

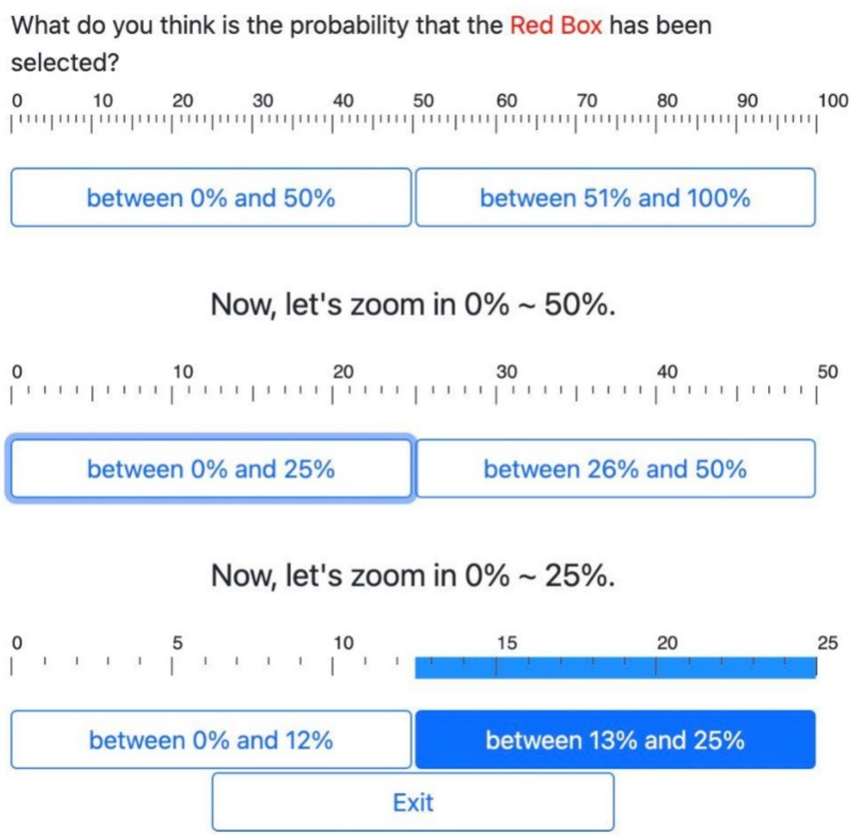
3.3 Experimental Design

In order to explore the empirical validity of DBM, we design the experiment with a collection of perception tasks that are used in the existing literature, and use the slider bar version of the CM as the benchmark, which allows both within-subject and between-subject investigations.

3.3.1 DBM and CM

We employ DBM to probe subjects' beliefs in a step-by-step manner. Initially, participants are queried about their assessment of probability relative to 50%. Subsequently, based on their response, they are prompted to determine whether the likelihood is below or above 25%, and this process continues iteratively. At each step, subjects are presented with two exclusive interval choices and the option to "Exit." Upon reaching the final step, participants must provide a point belief. Figure 3.1 shows the experimental interface of the DBM.

Figure 3.1: DBM: Experimental Interface



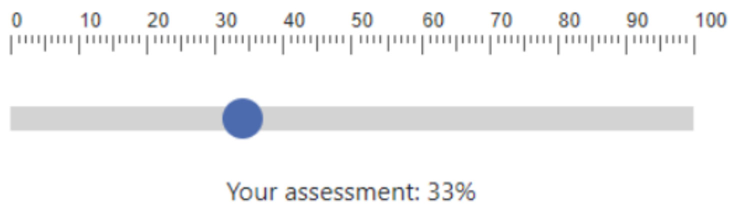
In instances where a subject provides a point belief, compensation is awarded based on the Binary Scoring Rule (BSR). Conversely, if an interval belief is reported, a random number within the specified interval is drawn from a uniform distribution. Subsequently, compensation is determined according to the BSR using the generated number.

We use the slider bar version of the CM to elicit subjects' probabilistic beliefs or perceptions, as it is widely used in many experimental studies.[3] To ensure accuracy and relevance to the task's objective truth, we offer three distinct scales: the 100 scale, 4000 scale, and 250 scale. Notably, we deliberately avoid providing a default position on the

[3]In the rest of the paper, we always refer CM as to the slider bar version of CM.

slider bar to mitigate any potential anchoring effects. Additionally, above the slider bar, we include a ruler for subjects' reference, aiding in their precise assessment. Figure 3.2 shows the experimental interface of the CM.

Figure 3.2: CM: Experimental Interface



3.3.2 Experiment 1: Within-subject Design

In Experiment 1, we use a within-subject design, allowing each subject to experience both the CM and DBM in a random order. To be more specific, the experiment consists of two blocks, with each block employing either CM or DBM to elicit subjects' beliefs or perceptions. Within each block, we use five different task categories regarding probabilistic beliefs or perceptions that are commonly used in existing literature. This design choice allows us to assess the generalizability of aggregate performances in the DBM compared to those in the CM.

Within each block, four of the five task categories use artificial settings such as balls and urns, peas, or dots, which are common in the laboratory experiments. These task categories include reporting prior belief [Danz et al., 2022], belief updating [Danz et al., 2022], forming compound prior belief [Liang, 2022], and estimating number of peas in a bowl [Falk and Zimmermann, 2018] or dots in a graph [Dewan and Neligh, 2020]. The fifth task category involves subjects' perceptions or probabilistic beliefs about real economic variables, specifically the inflation rate and the S&P500 [Enke and Graeber, 2023]. That is, there are two tasks within each task category: one from each block.

Each task has one objective truth, allowing us to measure the accuracy of subjects' beliefs or perceptions objectively. Detailed questions used in the experiment can be found in Appendix A.2. To prevent anchoring, we carefully choose parameters so that the objective truths in all tasks are spread across the entire range between 0 and 1.[4] To ensure comparability within the same task category, we deliberately choose parameters for tasks under the same category in the two blocks so that their objective truths are symmetric around 50%.[5] Table 3.1 demonstrates the five task categories and the corresponding objective truths used in the two blocks.

Table 3.1: Task Categories and Objective Truths in Experiment 1

	Task Category	Block 1	Block 2
4*Laboratory/Artificial Settings	Prior Belief	20%	80%
	Belief Updating	33%	67%
	Compound Prior	60%	40%
	Estimation	(peas) 3000/4000	(dots) 120/250
Real-life Settings	Econ Variable	(inflation rate) 92%	(S&P 500) 8%

Note: Within each block, the tasks and parameters are fixed, but the order of tasks is randomly determined.

Within each block, the tasks and parameters are fixed, but the order of tasks is randomly determined. We implement two treatments, Treatment DBM-CM and Treatment CN-DBM, by alternating the order of CM and DBM used to elicit beliefs or perceptions in the tasks of each block. That is, in Treatment DBM-CM, DBM is used to elicit subjects' beliefs or perceptions in Block 1, followed by CM in Block 2. Conversely, in Treatment CM-DBM, CM is used in Block 1, followed by DBM in Block 2. This design choice allows us to investigate the interaction between DBM, learning, and experience.

[4]For the estimation tasks involving counting peas in a bowl or dots in a graph, where the scales are $0-4000$ and $0-250$ respectively, we transform these into a $0-100$ scale to avoid duplicated objective truths with other probabilistic tasks.

[5]The objective truths in the estimation tasks involving counting peas in a bowl or dots in a graph do not have this property as the scales are different.

3.3.3 Experiment 2: Between-subject Design

The design of Experiment 1 could make differences across task categories difficult to interpret, as both the tasks and their objective truths vary. Subjects might differ in their expertise across task categories, and the measured accuracy may be influenced by different objective truths used. Empirical evidence demonstrates that individuals' subjective beliefs tend to be center-biased [Danz et al., 2022] or compressed towards an "intermediate" value, such as midpoint of a slider bar [Enke and Graeber, 2023]. Thus, even with the same subjective beliefs, performance may appear better when tasks have more centered objective truths (40% - 60%) compared to those with more extreme values (0% - 10%, or 90% - 100%) and those with intermediate values (10% - 40%, or 60% - 90%).

To address this concern, we design Experiment 2 with two main features: (1) within the same task category, we employ more objective truths that span the entire range between 0 and 1; and (2) we include three additional task categories with real-life settings alongside the existing belief updating tasks and perception tasks on inflation rate.

To be more specific, in Experiment 2, each subject needs to finish five tasks: two replicated from Experiment 1 (belief updating tasks and perception tasks about inflation rates) and three new perception tasks about real economic variables (income [Wiswall and Zafar, 2015a], unemployment rate, and education level).[6] By replicating tasks from Experiment 1, we can test the robustness of the results. Combining Block 1 of Experiment 1 and Experiment 2 provides a balanced set of tasks between laboratory/artificial settings and real-life settings and mitigates the learning effects. Each subject see the five tasks in a random order.

Within each task category, every subject randomly receives one of three parameters

[6]We generate questions about the unemployment rate and education level with objective truths using a method similar to the tasks on inflation rates in Enke and Graeber [2023].

corresponding to one of three types of objective truths: centered truths (40% - 60%), extreme truths (0% - 10% or 90% - 100%), and intermediate truths (10% - 40% or 60% - 90%), each of which is equally likely to occur. Table 3.2 depicts the task categories and objective truths used in Experiment 2. This design choice allows us to to distinguish the role of truth types from the impact of varied expertise across different task categories. We implement two treatments, Treatment CM and Treatment DBM, by using different methods to elicit beliefs or perceptions.

Table 3.2: Task Categories and Objective Truths in Experiment 2

	Task Category	Centered Truth	Intermediate Truth	Extreme Truth
4*Real-life Settings	Inflation Rate	56%	77%	92%
	Income	45%	30%	7%
	Unemployment Rate	56%	84%	98%
	Education Level	49	12	3
Laboratory/Artificial Settings	Belief Updating	47%	33%	6%

Note: The order of tasks is randomly determined.

3.3.4 Implementation and Recruitment Details

We recruited all subjects on Prolific, an online platform frequently used for research studies. To qualify for our study, subjects were required to have a minimum of 100 prior submissions on Prolific, with an approval rate of at least 98%. We implemented the experiment using the oTree platform [Chen et al., 2016]. For Experiment 1, we recruited 102 subjects, with 51 subjects assigned to each order of methods. For Experiment 2, we recruited 149 subjects, with 72 subjects using CM and 77 subjects using DBM to elicit beliefs in the five tasks. Each participant also received a $3 completion payment and took around 20 minutes to complete the study. In each experiment, subjects receive detailed instructions and are required to correctly answer comprehension questions before proceeding to the main parts of our study.

3.4 Results

We start by comparing the aggregate performance between DBM and CM using pooled data from Block 1 of Experiment 1 and Experiment 2, as shown in in Section 3.4.1. We find that DBM does not perform significantly different from CM at the aggregate level. Next, we investigate circumstances where DBM might outperform CM in Section 3.4.2. This includes examining whether the objective truth has extreme, centered or intermediate values, and whether the task context involves laboratory/artificial or real-life settings. We use pooled data from Block 1 of Experiment 1 and Experiment 2 to study these factors. We find that DBM outperforms CM when the objective truth is extreme, while CM outperforms DBM with intermediate objective truths. However, with centered truths and across task types, DBM does not perform differently from CM.

Then we analyze to what extent the length of stated beliefs in DBM informs the accuracy of subjects' beliefs in Section 3.4.3. Using pooled data from Block 1 of Experiment 1 and Experiment 2, we document that for stated interval beliefs in DBM ($a_l \neq a_u$), the shorter the interval, the more accurate the stated belief. However, stated point beliefs in DBM, which constitute a significant fraction of all stated beliefs, are not the most accurate. We find similar results using data from both blocks of Experiment 1 for within-subject analysis. Finally, we compare multiple methods for using stated beliefs in DBM to predict stated point beliefs in CM in Section 3.4.4 and discuss how to effectively utilize the data collected with DBM.

3.4.1 DBM vs. CM: Aggregate Performance

To test the empirical performance of DBM, we compare the accuracy of beliefs or perceptions elicited in DBM with those in CM. This requires a measure of accuracy. We mainly focus on two measures, given their stated belief $a = [a_l, a_u]$ and the objective truth q:

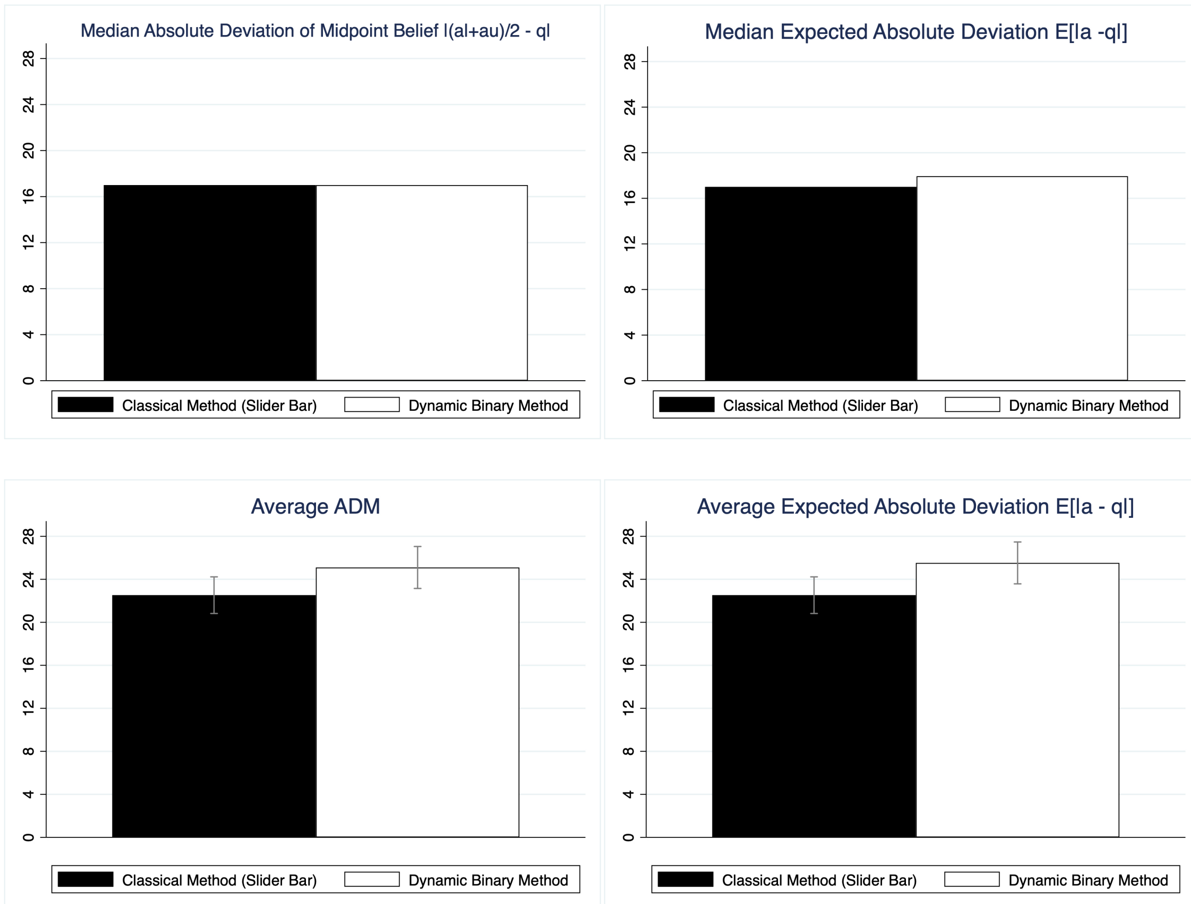
1. Absolute difference between the midpoint of their stated beliefs and the objective truth (ADM): $|\frac{a_l+a_u}{2} - q|$: ;

2. Expected absolute difference between their stated beliefs and the objective truth (EAD): $E_a[|a - q|]$.

Note that the use of midpoint in the first measure is justified by the theoretical framework that, for an expected utility maximizer, whether myopic or not, the midpoint of their stated belief reveals the mean of their true belief. Moreover, if the stated belief is a singleton, that is, $a_l = a_u$, the two measures are equivalent to $|a - q|$ – the absolute difference between stated belief and the objective truth.

Figure 3.3 demonstrates the median and mean accuracy of stated beliefs elicited with DBM versus CM using the two measures mentioned above. The median accuracy of stated beliefs elicited with DBM is not significantly different from those with CM, and this finding is robust across the measures used. Specifically, we conduct quantile regression of the measured accuracy on the dummy variable indicating which elicitation method is used (DBM or CM), controlling for gender and self-reported familiarity with statistics. The estimated coefficient on the elicitation method dummy variable is not significantly different from zero even at the 90% level ($p = 0.240$ for ADM and $p = 0.148$ for EAD).

Figure 3.3: Aggregate Accuracy of Stated Beliefs in DBM vs. CM



Note: Each graph uses pooled data from Block 1 of Experiment 1 and Experiment 2. For the bottom two graphs of average accuracy, we plot the 95% confidence intervals.

The mean accuracy of stated beliefs elicited using DBM is slightly lower than those using CM. We use OLS regression of the measured accuracy on the dummy variable indicating which elicitation method is used (DBM or CM), controlling for gender and self-reported familiarity with statistics.[7] The estimated coefficient on the elicitation method dummy variable is significantly different from zero at the 90% level: the average ADM using DBM is 2.88 larger than that using CM ($p = 0.098$), and the average EAD

[7]All the regression models have gender and self-reported familiarity with statistics controlled.

with DBM is 3.27 larger than that with CM ($p = 0.061$). This finding indicates that the aggregate performance of DBM is not significantly different from CM, although DBM exhibits slightly larger variance.[8]

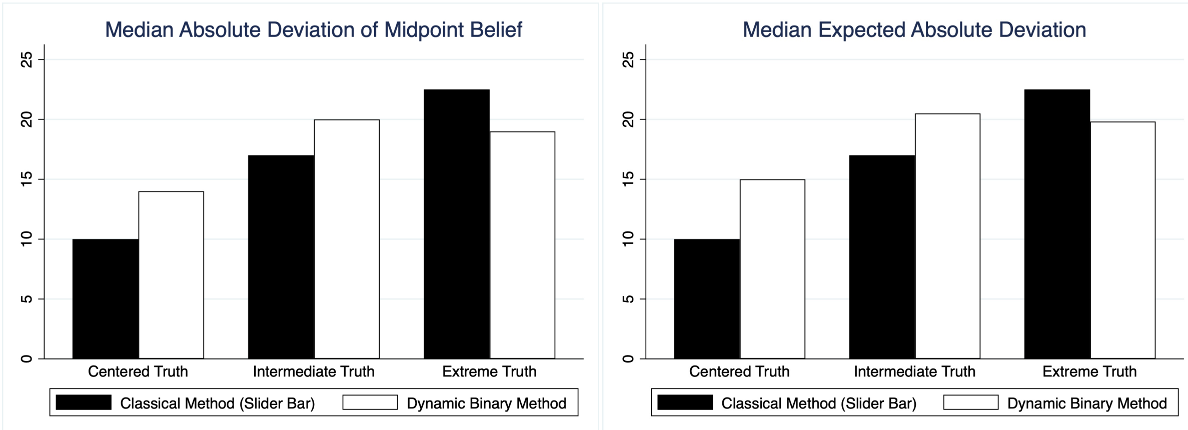
Result 4. *At the aggregate level, DBM does not perform significantly different from CM: the accuracy of stated beliefs elicited by the two methods are not significantly different.*

3.4.2 When does DBM Outperform CM?

Objective Truth Type. The null result at the aggregate level could be because DBM draws subjects' attention to non-centered values, thereby reducing subjects' tendency to choose numbers centered at the midpoint of the slider bar as their stated beliefs in each task. Figure 3.4 presents the median accuracy of beliefs elicited with DBM and CM using two measures, separated by the three types of objective truth: centered truths (40% - 60%), extreme truths (0% - 10% or 90% - 100%), and intermediate truths (10% - 40% or 60% - 90%).

[8] Similar to Enke and Graeber [2023]'s study and given that the directional results using average accuracy as the outcome variable are consistent with those using median accuracy but exhibit much larger variance, we primarily focus on median accuracy in the rest of the analysis to demonstrate the aggregate results of interest in the main draft.

Figure 3.4: Median Accuracy of Stated Beliefs in DBM vs. CM by Objective Truth Type



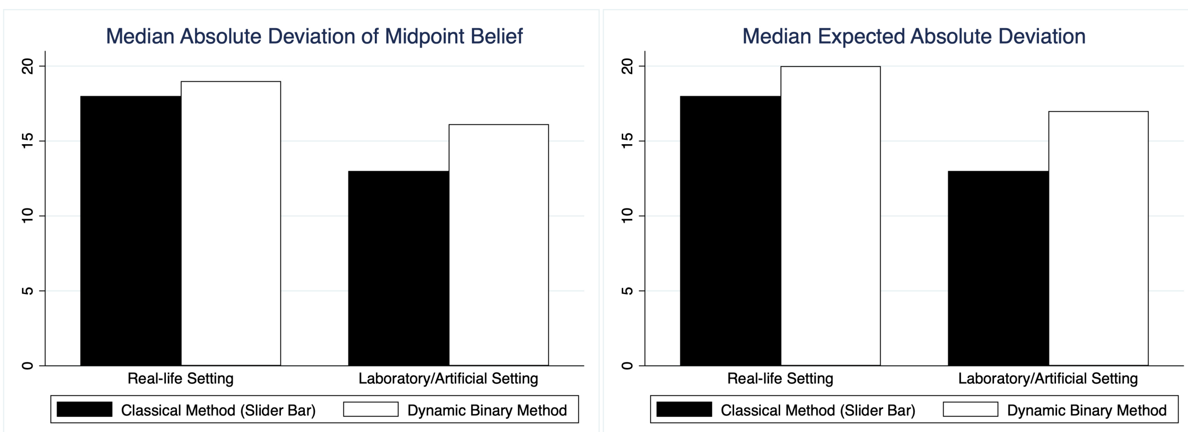
Note: Each graph uses pooled data from Block 1 of Experiment 1 and Experiment 2. Centered truths denote tasks with objective truths between 40% and 60%, extreme truths denote tasks with objective truths between 0% and 10% or between 90% and 100%, and intermediate truths denote tasks with objective truths between 10% and 40% or between 60% and 90%.

Consistent with Figure 3.4, the median accuracy of stated beliefs in DBM is significantly higher than in CM at the 95% confidence level when the objective truths are extreme (quantile regression, $p = 0.033$ for ADM and $p = 0.032$ for EAD). However, the median accuracy of stated beliefs in DBM is significantly lower than that in CM when the objective truths are intermediate (quantile regression, $p = 0.032$ for ADM and $p = 0.023$ for EAD). There are no significant differences for centered objective truths (quantile regression, $p = 0.134$ for ADM and EAD). This finding suggests that some deviations from the objective truths in CM could be attributed to the narrowed consideration set that subjects use to state beliefs or perceptions. Our new method, DBM, aids subjects by expanding the range of available numbers they consider.

Result 5. *DBM outperforms CM when the objective truth is extreme, while CM outperforms DBM with intermediate objective truths. However, there is no significant performance difference between DBM and CM when the objective truth is centered.*

Task Type. In addition to that, DBM may guide subjects to think through each task in a step-by-step manner, which could help retrieve information and past experiences from memory, especially for tasks with real-life settings that do not provide all the information needed for answering the task question correctly. Figure 3.5 demonstrates the median accuracy of beliefs elicited with DBM and CM using two measures, separated by the two types of tasks: tasks with real-life settings which involves subjects' perceptions of inflation rates, unemployment rates, income distribution, and education levels by state; and tasks with laboratory/artificial settings which includes those on prior beliefs, belief updating, counting, and compound priors.

Figure 3.5: Median Accuracy of Stated Beliefs in DBM vs. CM by Task Type



Note: Each graph uses pooled data from Block 1 of Experiment 1 and Experiment 2. Tasks with laboratory/artificial settings include those on prior beliefs, belief updating, counting, and compound priors. Tasks with real-life settings involve subjects' perceptions of inflation rates, unemployment rates, income distribution, and education levels by state.

As shown in Figure 3.5, the median accuracy of beliefs elicited using DBM is not significantly different from CM in tasks with real-life settings (quantile regression, $p = 0.489$ for ADM and $p = 0.484$ for EAD).[9] In tasks with laboratory/artificial settings,

[9] As the results with ADM are similar to those with EAD, we primarily use EAD as the measure of accuracy in the remainder of the analysis.

the median accuracy using DBM is slightly lower than CM, but the significance of this result depends on the measure of accuracy (quantile regression, $p = 0.126$ for ADM and $p = 0.038$ for EAD). This indicates that DBM does not perform differently from CM across task types.

Result 6. *DBM does not perform differently from CM regardless of whether the task utilizes a laboratory/artificial setting or a real-life setting.*

3.4.3 Is the Interval Length Informative About Accuracy?

This section explores the relationship between the length and the accuracy of stated beliefs in the tasks using DBM. To ensure the lengths of stated beliefs are comparable across tasks, for all the analysis in this subsection, we use only tasks with a choice scale of 100, which rules out tasks with counting peas in a bowl and counting dots in a graph in Experiment 1. To achieve this goal, we start with pooled data from Block 1 of Experiment 1 and Experiment 2 to conduct a between-subject analysis, investigating how the median accuracy of stated beliefs in tasks using DBM varies with the number of steps taken. Additionally, we use the data from Blocks 1 and 2 in Experiment 1 to explore, from a within-subject perspective, to what extent the length of a subject's stated belief in tasks using DBM can predict how well they perform in tasks using CM.

Our theoretical framework shows that an exit early before reaching the point indicates that the expected utility maximizer is myopic — failing to perfectly foresee that choosing until $a^* = E(p)$ is optimal, regardless of the precision of their true beliefs. Conversely, choosing until the end suggests that the expected utility maximizer is not. If this is the case, we would expect no correlation between the length of stated beliefs (number of steps taken) and their accuracy.

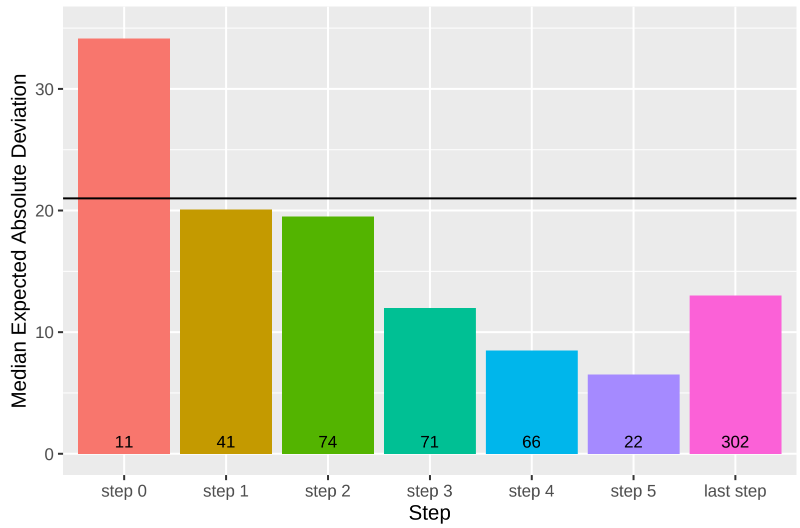
Figure 3.6 plots the median EAD of stated beliefs using DBM against the number of steps taken. Generally, the median EAD of stated beliefs decreases as the number of steps increases, with the correlation being significantly different from zero at the 95% confidence level (quantile regression, $p = 0.018$). However, for the stated point beliefs in DBM, which constitute 51% of all the stated beliefs, the median EAD is slightly higher than those exiting right before the last step (i.e., Step 5). Similar patterns are observed when separated by task types and by objective truth types, as shown in Figure 3.7. This finding could result from overconfidence – where individuals overestimate the precision of their perceptions – or from risk aversion – where individuals dislike uncertainty in their reported beliefs. Distinguishing between potential mechanisms could be a fruitful direction for future research.

Result 7. *The length of stated beliefs in DBM is negatively correlated with their own accuracy at the aggregate level: the longer the interval, the less accurate the stated belief. However, this relationship is not strictly monotonic: stated beliefs reaching the point are not the most accurate ones.*

Moreover, we use data from Blocks 1 and 2 in Experiment 1 to study this question from a within-subject perspective. Specifically, since tasks from the same category between blocks have symmetric objective truths, we pair tasks from the same task category – one using DBM to elicit beliefs and the other using CM. Within each pair, we investigate the extent to which the length of stated belief in the task using DBM can predict the accuracy of stated belief in the the task using CM.

Figure 3.8 plot the median absolute deviation in tasks using CM against the number of steps taken in the paired tasks using DBM. When pooling Treatments DBM-CM and CM-DBM, there is no significant correlation between the median absolute deviation of stated beliefs with CM and the number of steps taken in their paired tasks with DBM. This

Figure 3.6: DBM: Median Accuracy of Stated Beliefs and Number of Steps Taken

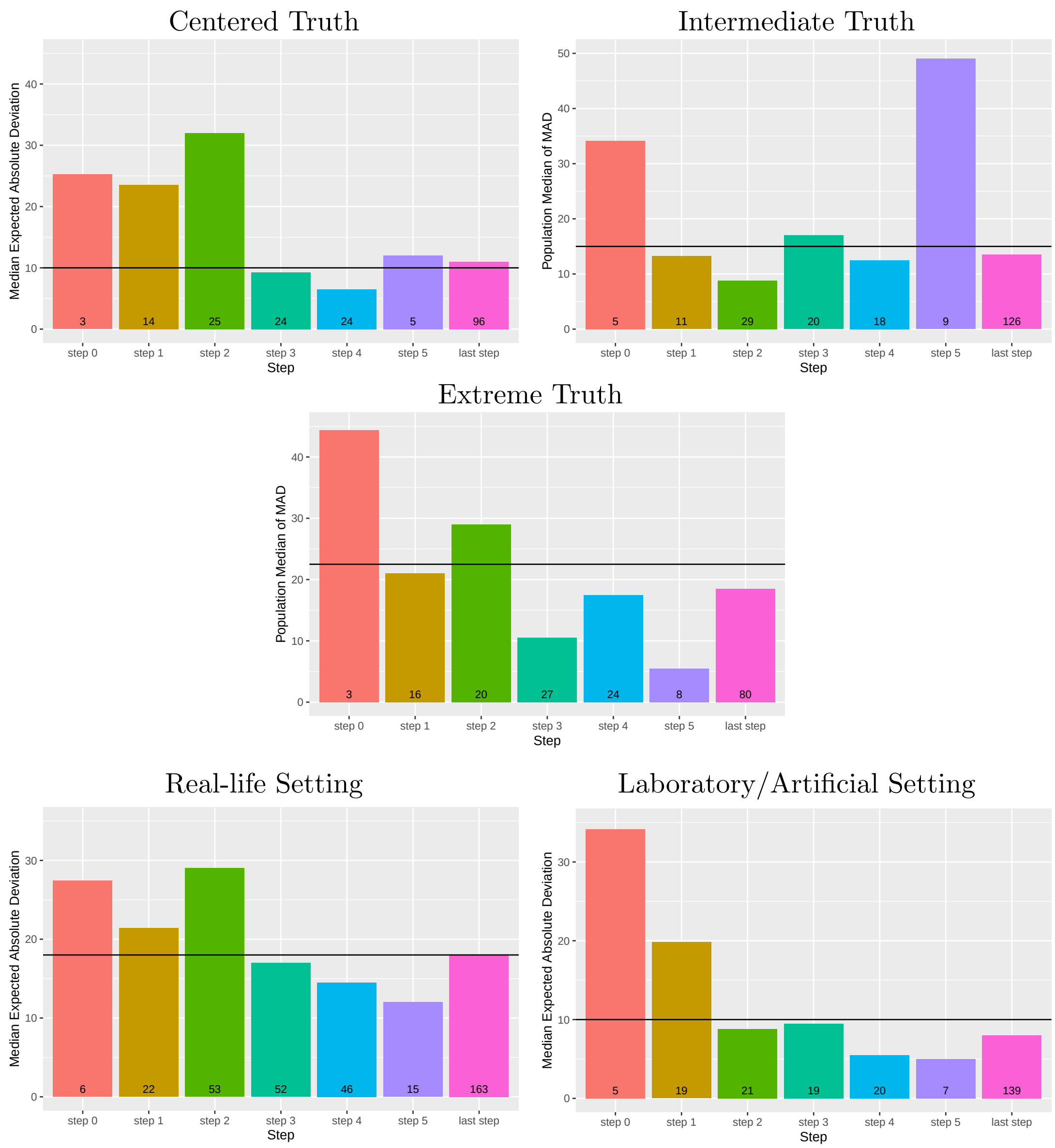


Note: The number labeled inside each bar is the number of stated beliefs that exit in each step. The black horizontal line is located at the median expected absolute deviation with all the stated beliefs using CM pooled.

mainly results from the null effect in Treatment CM-DBM.[10] When separating the data by treatment, we find that in Treatment DBM-CM – using DBM in Block 1 and CM in Block 2 – the median absolute deviation of stated beliefs with CM significantly decreases as the length of stated beliefs increases (quantile regression, $p > 0.001$). Similar to the between-subject analysis discussed earlier, the relationship is not strictly monotonic: point or close-to-point beliefs in tasks using DBM do not predict the lowest median absolute deviation in paired tasks using CM.

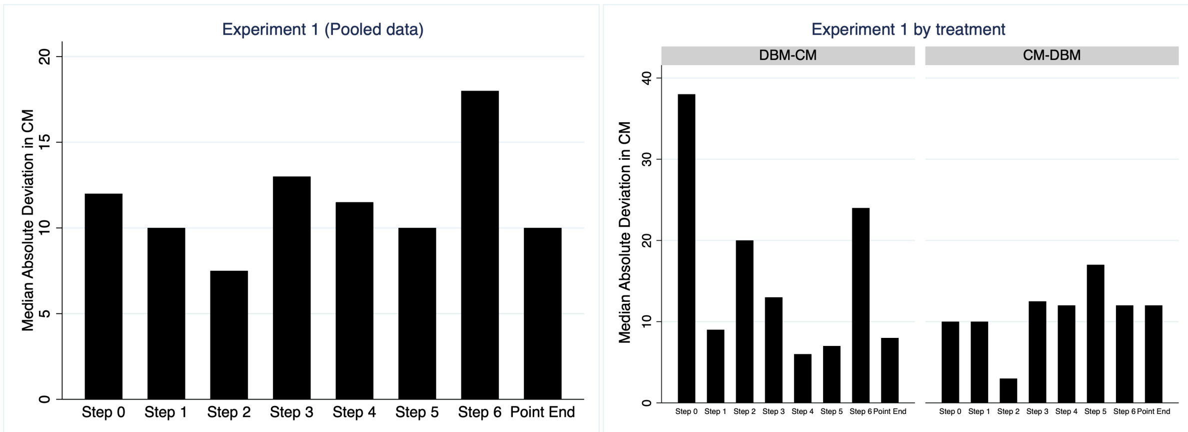
[10] One reason the data in Treatment DBM-CM are much noisier is that subjects took longer than expected to complete Experiment 1, resulting in a base payment that was considered as lower than the recommended hourly rate by Prolific. Since DBM requires more time for subjects to think through and submit their beliefs, the quality of choices decreases due to fatigue when DBM is used in Block 2.

Figure 3.7: DBM: Accuracy and Number of Steps Taken by Objective Truth Type and Task Type



Note: The number at the bottom of the bar is the number of observations. The black horizontal line is the median absolute deviation in CM.

Figure 3.8: Median Accuracy of Stated Beliefs in CM and Number of Steps Taken in DBM

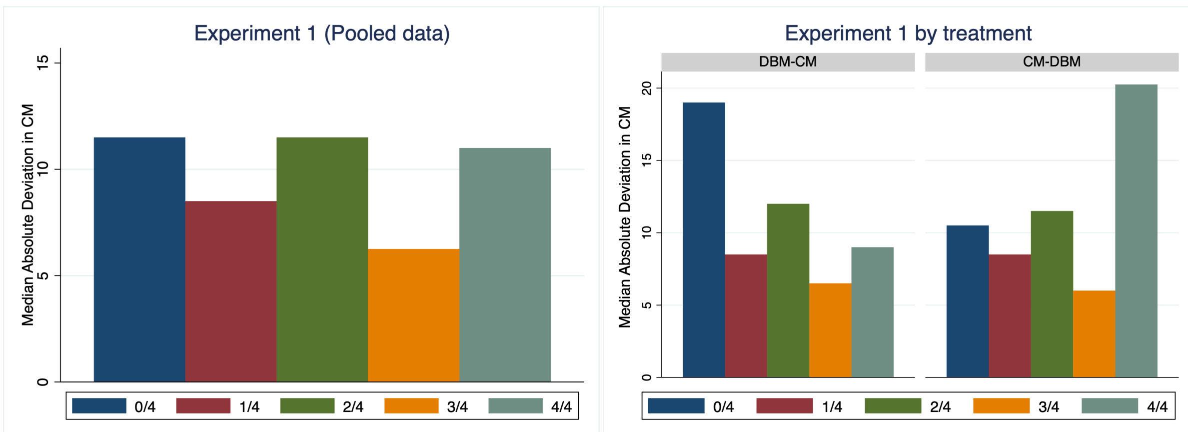


Note: X-axis denotes the number of steps taken by the stated beliefs in DBM. Each graph uses pooled data from Experiment 1. The left panel pooled data from Treatments DBM-CM and CM-DBM together, while the right panel separates results by treatment.

A similar pattern is observed in Figure 3.9, where we calculate, for each subject, the median accuracy of stated beliefs among the four paired tasks using CM, separated by the number of point beliefs they stated among the four tasks using DBM. Using pooled data, among those who state interval beliefs in at least one of the four tasks using DBM, the median subject who is more likely to state point beliefs in tasks using DBM is also more accurate in tasks using CM (quantile regression, $p = 0.015$). However, for subjects who state point beliefs in all the four tasks using DBM, the median subject is not the most accurate one in tasks using CM. This pattern is even stronger in Treatment DBM-CM (quantile regression, $p = 0.03$).

Result 8. *The length of stated beliefs in tasks using DBM is negatively correlated with the accuracy of stated beliefs in tasks using CM: the longer the interval in DBM, the less accurate the stated belief in CM. However, this relationship is not strictly monotonic, as stated beliefs reaching the point in DBM do not predict the most accurate beliefs stated* *in CM.*

Figure 3.9: Within Subject: Median Accuracy in CM and Fraction of Point Beliefs in DBM



Note: X-axis denotes the fraction of point beliefs in DBM. Each graph uses pooled data from Experiment 1. The left panel pooled data from Treatments DBM-CM and CM-DBM together, while the right panel separates results by treatment.

3.4.4 Predicting Point Beliefs Elicited with CM

In this section, we explore multiple methods for using the stated beliefs elicited with DBM to predict point beliefs elicited with CM. To achieve this, we primarily focus on data from Experiment 1 with a scale of 100. We use the paired stated beliefs in tasks using DBM to predict the corresponding beliefs in tasks using CM. To ensure comparability, we symmetrize the stated belief and objective truth for one task in each pair. There are several methods to utilize the data elicited with DBM:

1. Midpoint Prediction: according to our theoretical framework, the midpoint of stated beliefs using DBM serves as a natural predictor for stated point beliefs in CM:

$$\hat{a}_1 = \frac{a_l + a_u}{2}$$

where a_l and a_u denote the lower and upper bound of stated beliefs in DBM. For an expected utility maximizer, they will select the mean of their true belief as stated belief in CM and will choose the belief whose midpoint is equal to the mean of their true belief in DBM.

2. Cognitive Default, Cognitive Noise, and Objective Truth: Existing studies indicate that individuals' stated beliefs in CM often compress towards a cognitive default (e.g., the center of the slider bar) [Danz et al., 2022, Enke and Graeber, 2023]. This phenomenon can be modeled as a weighted average between the utility-maximizing decision $a^*(p)$ and the cognitive default d. The relative weight on d is determined by the cognitive noise or uncertainty λ [Enke and Graeber, 2023]:

$$\hat{a}_2 = (1 - \lambda) * a^*(p) + \lambda * d$$

The greater the cognitive noise, the stronger the tendency to state the default d (e.g., center of the slider bar) in CM. To construct the predicted beliefs in CM, we use a straightforward method to determine λ: $\lambda = \frac{a_u - a_l}{100}$, where λ is the length of stated beliefs in the paired tasks using DBM relative to the scale. Using the objective truth $p = a^*(p)$ and the default $d = 50\%$, we can generate the predicted point beliefs in CM.

3. Cognitive Default, Cognitive Noise, and Subjective Truth: We propose a revised version of Method 2 by replacing the objective truth, which could be equally difficult for subjects with bounded rationality to perceive, with the subjective truth – the midpoint of stated beliefs in DBM:

$$\hat{a}_3 = (1 - \lambda) * \frac{a_l + a_u}{2} + \lambda * d$$

where $\lambda = \frac{a_u - a_l}{100}$ and the default $d = 50\%$.

Table 3.3 show the average predicted point beliefs using stated beliefs in DBM via the three methods mentioned above, separated by task category. Using t-tests to contrast with the average stated beliefs in CM, we find that the predicted beliefs using Method 3 – weighted average between cognitive default and subjective truth – are closest to the stated beliefs in CM. Specifically, using Method 3, the predicted beliefs in tasks with economic variables are not significantly different from the average stated beliefs in CM. Additionally, using Method 3, the predicted beliefs in tasks about Simple Prior and Compound Prior differ from those in CM at a 95% confidence level, which is a much smaller difference compared to the other two methods.

Table 3.3: Predicted Point Beliefs using Data in DBM and Stated Point Beliefs in CM

Task Category	Stated Beliefs in CM		Method 1		Method 2		Method 3
Simple Prior	30.3%	**	35.1%	***	22.2%	**	36.1%
Compound Prior	54.9%	***	61.9%	***	59.3%	**	61%
Posterior	43.1%	***	51.1%	***	35.2%	***	51%
Econ Variables	67.3%		70.8%	***	87.1%		69.9%

Note: Reported significance stars are based on a two-way t-test to determine whether the difference between the average stated beliefs in CM and the average predicted beliefs using each of the three methods is significantly different from zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Conversely, the predicted point beliefs using Method 2 – weighted average between cognitive default and subjective truth – are significantly different from those in CM at the 99% confidence level in each of the task categories. Method 1, which uses the midpoint, performs somewhere in between the other two methods. One plausible reason for the worst performance of Method 2 is that its predictions are constrained between the objective truth and the default, failing to capture stated beliefs in CM that fall outside this range.

3.5 Conclusion

In this paper, we propose a novel method, the *Dynamic Binary Method* (DBM), to elicit people's beliefs or perceptions. Beliefs and perceptions are central to studying economic behavior, yet accurately eliciting them presents significant challenges. Existing elicitation methods that involve soliciting respondents' best point estimates are susceptible to various biases, such as the cognitive difficulty of pinpointing imprecise beliefs. Unlike *Classical Methods* (CM), which require absolute judgments on the best point estimates of beliefs, DBM, inspired by the bisection method, prompts respondents to make a series of binary relative judgments. This approach allows respondents to state interval beliefs by exiting the process at any step before reaching a final point estimate.

To assess the empirical validity of DBM, we use a collection of perception tasks from existing literature, construct both between-subject and within-subject experiments, and use the slider bar version of CM to benchmark how well respondents would perform in each task.

The main finding is that, at the aggregate level, DBM does not perform significantly differently from CM, regardless of whether the perception question uses artificial/laboratory settings or real-life settings. This finding is robust to different measures we use, including the absolute deviation of the midpoint of elicited beliefs from the objective truth or the expected absolute deviation of elicited beliefs from the objective truth. However, DBM outperforms CM when the task has extreme values as the objective truth. This suggests that some biases in perception questions could result from the narrowed consideration set that respondents use to choose beliefs or perceptions from.

Furthermore, we find, from both between-subject and within-subject perspectives, that the length of stated beliefs in tasks using DBM is negatively correlated with their own accuracy and can predict how well respondents perform in tasks using CM at the aggregate level: the longer the interval, the less accurate the stated belief in DBM and the less accurate the stated belief in CM. However, this relationship is not strictly monotonic: stated beliefs reaching the point are not the most accurate ones and do not predict the most accurate beliefs stated in CM.

Lastly, we compare three methods of using the stated beliefs elicited with DBM to predict point beliefs elicited with CM. We find that predictions using a weighted average between subjective truth (the midpoint of stated beliefs in DBM) and the cognitive default (e.g., midpoint of the slider bar), with the relative weight on the default determined by the length of stated beliefs in DBM, are closest to the average stated beliefs in CM. This approach outperforms both using the midpoint of stated beliefs in DBM alone and using objective truth instead of subjective truth in the weighted average method. Our findings underscore the significance of incorporating the precision of stated beliefs and perceived truth to enhance predictions of economic behavior.

Our results also raise several intriguing questions for future research. As explored in the literature review, there are non-incentivized methods for identifying preference incompleteness, taste imprecision, or the distribution of beliefs. It would be valuable to compare these methodologies to understand the degree to which they capture the same uncertainty in beliefs or perceptions and how this may differ between preference incompleteness and belief imprecision.

Additionally, it would be beneficial to gain a deeper understanding of the circumstances under which individuals possess precise versus imprecise beliefs. Such insights could aid in interpreting standard belief data and potentially enable the identification of imprecision even when individuals are unable to directly report it. Furthermore, it would be fascinating to explore any neurological or biological indicators of imprecision, which could provide a new dimension to understanding how individuals form and report their beliefs or perceptions.

Appendix A

Interfaces

A.1 Interfaces for RT Project

PART I

Part 1 consists of 10 rounds, in which each of you will play the role of a sender. In each round, the computer generates outcomes of 6 dice. The outcome of each die is randomly drawn from the set 1, 2, 3, 4, 5, 6.

We define the outcomes 1, 2, 3 as "small numbers", and 4, 5, 6 as **"large numbers"** .

You can privately learn the outcomes of the 6 dice by clicking the dice buttons on your screen. Each button will uncover the outcome of 1 die, and you can uncover as many as you want. The outcomes of the dice will not change with the total buttons you click to uncover. Then you need to report on the question if there are 4 or more "large numbers" by choosing "Yes" or "No". Your report could be independent from the outcomes of the 6 dice.

Your payment as a sender

Your payment will be based on the Parts 2 and 3 of the experiment. You will be matched with another player, who's a receiver, and his choices will determine your payoff as a sender.

The receiver in Parts 2 and 3 has to answer the same question as in Part 1, without the ability to learn the true outcomes. However, he will receive your report to inform him to guess the answer.

If the receiver chooses **Yes** (there ARE 4 or more "large numbers"), you will get **$8** , regardless of the true answer.

If the receiver chooses **No** (there ARE **NOT** 4 or more "large numbers"), you will get **$4** , regardless of the true answer.

The receiver gets $ 8 if his guess is right, and $4 if his guess is wrong.

If the instructions is clear to you, you may go to the next screen to answer comprehension questions.

OK

You will be able to proceed after answering the following questions correctly.

1. If the outcomes of 6 dice are 5,3,3,6,4,1, which statement is true?

A. It is true that there are 4 or more large numbers.

B. It is wrong that there are 4 or more "large numbers".

C. A sender has an obligation to report "No" based on this outcome.

- A
- B
- C

2. How is the sender's payment determined?

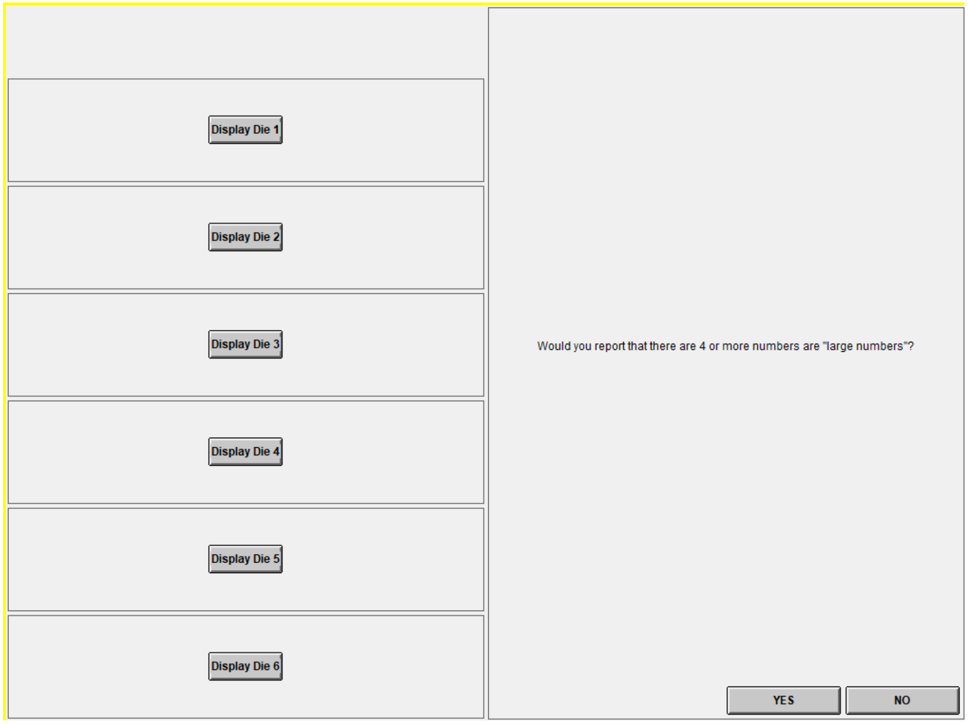
A. It depends whether the sender answers the questions correctly, if it is correct, the sender will get $8, otherwise she will get $4.

B. Only the receiver's choice will determine the sender' payoff. If the receiver chooses Yes, then the sender will get $8. If the receiver chooses No, then the sender will get $4.

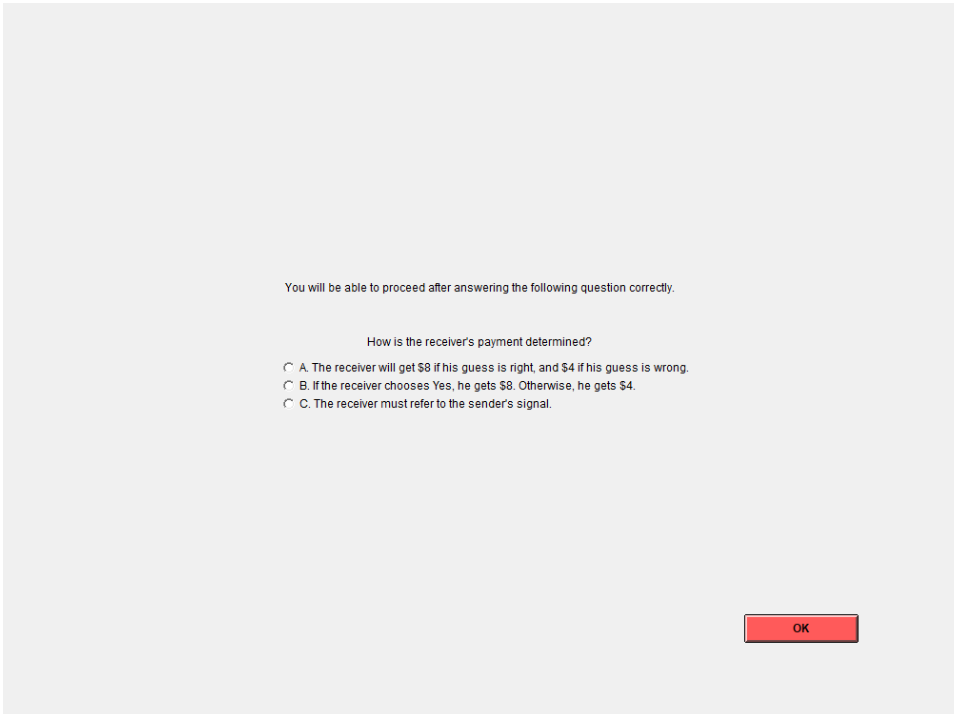
C. The sender only sends message to the receiver, the receiver's choice is going to determine the sender' payoff. If the receiver chooses correctly, then the sender will get $8. If the receiver chooses wrongly, then the sender will get $4

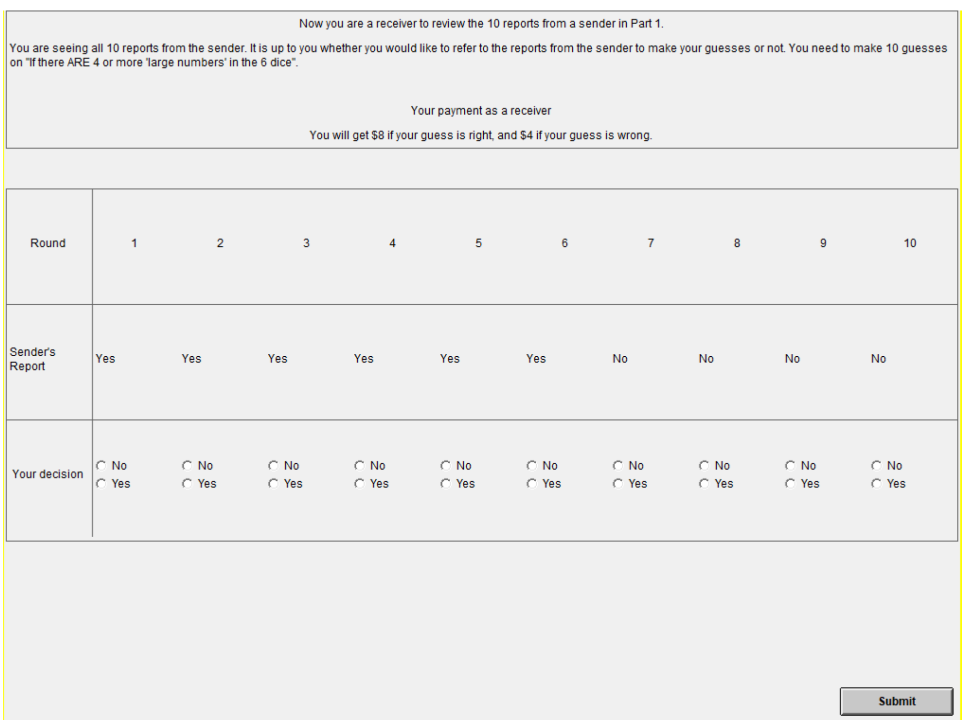
- A
- B
- C.

OK


Display Die 1
Display Die 2
Display Die 3
Display Die 4
Display Die 5
Display Die 6
Would you report that there are 4 or more numbers are "large numbers"?
YES
NO


PART II
Now you are a **receiver, who will get 10 reports from a sender in Part 1, one for each round. You'll get reports all at the same time.**
It is up to you whether you would like to refer to the reports from the sender to make your guesses or not. You will need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice". Your guess will be either **"Yes"** (there ARE 4 or more "large numbers") or **"No"** (there ARE NOT 4 or more "large numbers").
Your payment as a receiver
For any question, you will get $8 if your guess is right, and $4 if your guess is wrong.
If the instructions is clear to you, you may go to the next screen to answer comprehension questions.
OK


You will be able to proceed after answering the following question correctly.
How is the receiver's payment determined?
A. The receiver will get $8 if his guess is right, and $4 if his guess is wrong.
B. If the receiver chooses Yes, he gets $8. Otherwise, he gets $4.
C. The receiver must refer to the sender's signal.
OK


Now you are a receiver to review the 10 reports from a sender in Part 1.
You are seeing all 10 reports from the sender. It is up to you whether you would like to refer to the reports from the sender to make your guesses or not. You need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice".
Your payment as a receiver
You will get $8 if your guess is right, and $4 if your guess is wrong.
Round 1 2 3 4 5 6 7 8 9 10
Sender's Report Yes Yes Yes Yes Yes Yes No No No No
Your decision No Yes No Yes No Yes No Yes No Yes No Yes No Yes No Yes No Yes No Yes
Submit

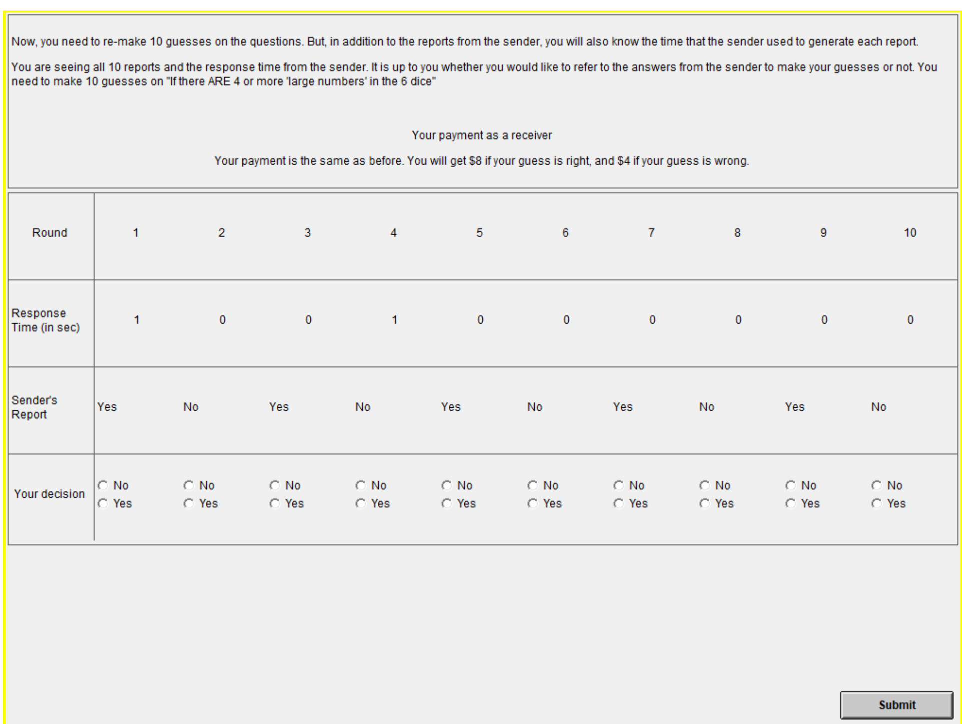
PART III

Now, you need to re-make 10 guesses on the questions. But, in addition to the reports from the sender, you will also know the time that the sender used to generate each report.

Your payment as a receiver

Your payment is the same as before. For any question, you will get $8 if your guess is right, and $4 if your guess is wrong.

OK



Now, you need to re-make 10 guesses on the questions. But, in addition to the reports from the sender, you will also know the time that the sender used to generate each report.

You are seeing all 10 reports and the response time from the sender. It is up to you whether you would like to refer to the answers from the sender to make your guesses or not. You need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice"

Your payment as a receiver

Your payment is the same as before. You will get $8 if your guess is right, and $4 if your guess is wrong.

Round	1	2	3	4	5	6	7	8	9	10
Response Time (in sec)	1	0	0	1	0	0	0	0	0	0
Sender's Report	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Your decision	No Yes	No Yes	No Yes	No Yes	No Yes	No Yes	No Yes	No Yes	No Yes	No Yes

Submit

Survey

By completing this survey, you will receive additional $1, which will be added to your total earnings.

1. What's your gender?
- Man
- Woman
- Other

2. Please indicate which of the following adjectives best describe yourself. Check all that apply.

Capable	Egotistical	Conventional
Honest	Original	Self-confident
Artificial	Commonplace	Informal
Intelligent	Narrow interests	Sexy
Clever	Humorous	Dissatisfied
Well-mannered	Reflective	Submissive
Cautious	Conservative	Insightful
Wide interests	Sincere	Snobbish
Confident	Individualistic	Suspicious
Inventive	Resourceful	Unconventional

OK

A.2 Questions used in Experiments 1 and 2

A.2.1 Experiment 1 Question Examples

1. **Inflation Rate** The computer randomly picked a year X between 1980 and 2018.

   What do you think is the chance that the U.S. inflation rate in year X was lower than 7.4%?

   In other words, imagine that, at the beginning of Year X, the set of products that is used to compute the inflation rate cost \$100. What do you think is the chance that, at the end of that same year, the same set of products cost less than \$107.4?

2. **S&P 500** The S&P 500 is an American stock market index that includes 500 of the largest companies based in the United States.

   The computer randomly picked a year X between 1980 and 2018.

   What do you think is the chance that the annual change rate of S&P 500 in Year X is less than -13%, i.e., the S&P 500 lost more than 13% of its value?

   In other words, imagine that someone invested \$100 into the S&P 500 at the beginning of Year X. What do you think is the chance that, at the end of that same year, the value of the investment was less than \$87?

3. **Prior Probability** What do you think is the likelihood (percent chance) that the selected box is the Red box, the one with more red balls? (round to the nearest integer)

4. **Posterior Probability** To give you a hint of which box was selected, the computer drew a ball from the selected box.

   The drawn ball is red. What do you think is the likelihood (percent chance) that the selected box is the Red box, the one with more red balls? (round to the nearest integer)

5. **Compound Lottery** This is either a 30-70 race or a 90-10 race.

   There is a 50% chance that this is a 30-70 race, otherwise this is a 90-10 race.

   What do you think is the chance that the Red horse won?(round to the nearest integer)

6. **Count Peas** How many peas are there in the bowl in the picture on the left? (round to the nearest integer)

7. **Count Dots** How many dots are there in the picture on the left? (round to the nearest integer)

A.2.2 Experiment 2 Question Examples

1. **Income** In 2022, among all individuals aged 30, what is the percentage of those that are working full time that earn $125,000 and above per year?

2. **Inflation Rate** Randomly pick a year from 1980 - 2018, what is the chance that the inflation rate in that year is lower than 2.6%?.

   In other words, imagine that, at the beginning of Year X, the set of products that is used to compute the inflation rate cost $100. What do you think is the chance that, at the end of that same year, the same set of products cost less than $102.6?

3. **Education** How many states in 2019 have less than 29% of state-level population that have Bachelor's degree or higher?

4. **Unemployment Rate** Randomly pick a year from 1980 - 2022, what is the chance that the unemployment rate in that year is lower than 5.5%?

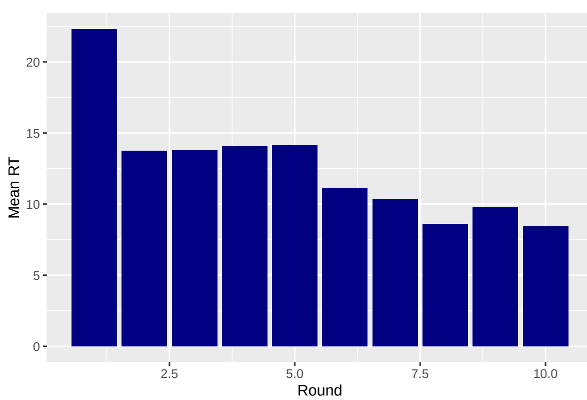
5. **Posterior Probability** 50%, 50% priors, 15:1 red/blue balls in the Left urn, 1:17 red/blue in the Right urn, random draw one that is red, What's the probability it comes from the Left urn?

Appendix B

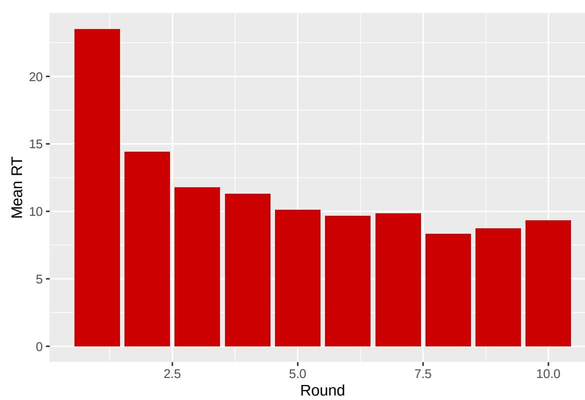
Graphs

B.1 Time trend of RT

In both conditions, round 1 took much more RT than the following ones, suggesting that subjects used it as a practice. In round 2-10, the latter rounds took less time than the former rounds.



(a) SU condition



(b) SA condition

Figure B.1: Mean of RT in each round.

B.2 Fraction of number of buttons clicked for different messages

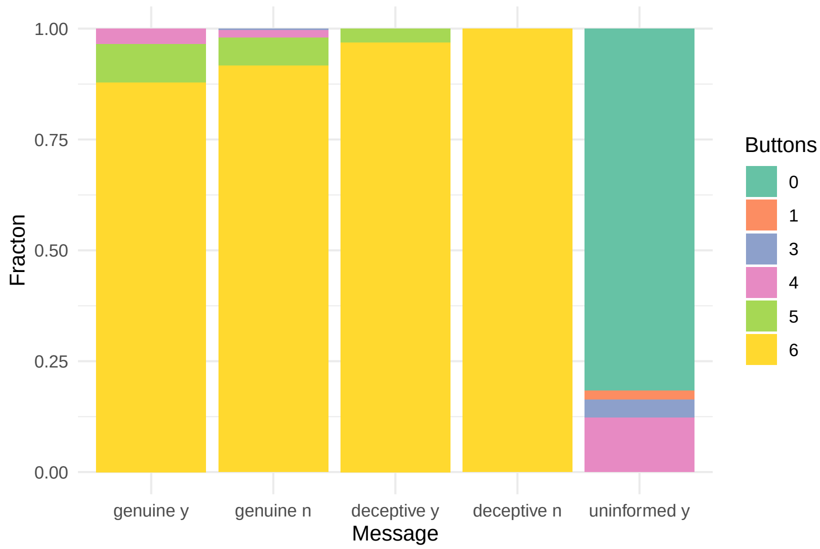


Figure B.2: Fraction of Different Numbers of Button Clicks. Different color represent different number of button clicks for each report.

Appendix C

Statistics

C.1 Payoff Comparisons

Table C.1 and C.2 shows that there are no significant differences in senders' and receivers' payoffs between RU and RI conditions. The difference between the genuine y and deceptive y suggests that receivers could somehow identify the deceptive y. No difference between the RU and the RI conditions suggests that receivers did not use the RT to identify the deceptive y. We find that the difference disappears when the count of y reports falls below than 7. It suggests that receivers use the total number of y reports from the sender to identify the deceptive y, and they did not trust the y from a sender if she sent too many y reports.

Message (obs.)	RU	RI	t test
genuine y (86)	\$7.35	\$7.16	p=0.43
deceptive y (51)	\$6.27	\$6.35	p=0.84
genuine n (138)	\$4.46	\$4.38	p=0.56
deceptive n (1)	\$4	\$4	not applicable
uninformed y (24)	\$4.17	\$4.17	p=1

(a) In the SU condition

Message (obs.)	RU	RI	t test
genuine y (87)	\$7.31	\$7.45	p=0.53
deceptive y (160)	\$6.58	\$6.84	p=0.50
genuine n (45)	\$4.35	\$4.50	p=0.28
deceptive n (3)	\$5.33	\$4	p=0.42
uninformed y (25)	\$6.72	\$6.24	p=0.39

(b) in the SA condition

Table C.1: Senders' Payoff Comparison

Message (obs.)	RU	RI	t test
y (161)	\$6.78	\$6.67	p=0.55
n (139)	\$7.51	\$7.60	p=0.57

(a) In the SU condition

Message (obs.)	RU	RI	t test
y (157)	\$6.60	\$6.52	p=0.73
n (163)	\$7.61	\$7.44	p=0.23

(b) In the SA condition

Table C.2: Receivers' Payoff Comparison

C.2 Regression

For the robustness check of *Result 5*, we employ an fixed effect OLS regression model as framework C.1 for 4 subsets, including $m = n$ and trusted without RT, $m = n$ and distrusted without RT, $m = y$ and trusted without RT, $m = y$ and distrusted without RT:

$$change_{ir} = a_i + b \times RTscale_{ir} + c_r + \epsilon_{ir} \tag{C.1}$$

The variable of interest is $change_{ir}$, which takes a value of 1 if the receiver i changed her guess at round r , and 0 if receiver did not change her guess. $RTscale_{ir}$ assumes a crucial role by effectively categorizing receiver i's observed RT at round r into three distinct classifications: Median of 10 observed RT, Below Median, or Above Median. We take the Below median as the benchmark, and b capture the effect difference between Median, Above and Below. We also control individual fixed effect and round fixed effect by a_i and c_r. ϵ_{ir} captures the random error. Table C.3 presents the outcomes of the fixed-effect OLS regression.

In result Column (1)-(4), we present the findings pertaining to the situation under the SU condition. Column (1) states that when receivers change from trust some high-payoff message y to not trust, they did not rely on RT. Column (2) states that receivers were 28% more likely to trust a longer y message than a short one. Column (3) states that receivers were 9% more likely to distrust a long n message than a short message. There are only 16 observations in Column (4), we prefer not to infer too much information from it. Overall, when receives had better trust relative short y message, they did trust long y message; when receivers' should always trust n message, their trust did rely somehow on RT, and they trust the shorter one more than the long one. This finding suggests that receivers are not clear about the relationship between RT and veracity of the message in the SU condition. They did not show consistent suspicion toward long or short messages.

	Dependent variable: change							
	SU				SA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	y,T	y,NT	n,T	n,NT	y,T	y,NT	n,T	n,NT
Median	-0.002	0.281*	0.080	−2.000***	−0.191*	0.192	0.040	1.273
	(0.136)	(0.149)	(0.086)	(0.000)	(0.104)	(0.329)	(0.073)	(0.997)
Above	-0.045	0.288***	0.089*	0.000	-0.132	0.115	-0.028	0.455
	(0.070)	(0.102)	(0.049)	(0.000)	(0.083)	(0.157)	(0.051)	(0.745)
Individual FE	✓	✓	✓	✓	✓	✓	✓	✓
Round FE	✓	✓	✓	✓	✓	✓	✓	✓
Constant	0.215	-0.137	-0.035	0.000	0.336	-0.188	-0.052	0.364
	(0.176)	(0.172)	(0.119)	(0.000)	(0.228)	(0.270)	(0.126)	(0.873)
Observations	102	59	123	16	118	39	148	15
R^2	0.485	0.586	0.518	1.000	0.399	0.788	0.454	0.708
Adjusted R^2	0.200	0.385	0.331	1.000	0.110	0.575	0.264	-0.364
Note:					*p<0.1; **p<0.05; ***p<0.01			

Table C.3: The Effect of RT on Receiver's Change Decision

They sometimes suspected the short one, sometimes suspected the long one. As a result, even though they tried to use it, they used it ineffectively.

Column (5)-(8) present the findings in the SA condition. The overall result says that receivers seldom used RT to make decisions in this condition, which was pretty rational given that senders successfully manipulated their RT. We present our result of receivers' behavior.

Appendix D

Proofs

D.1 Binarized Scoring Rule and Incentive Compatibility

Consider a decision maker (DM) with a probabilistic belief over a verifiable binary outcome $s \in \{A,\ B\}$, assuming they possess a true belief $p = Pr\{s = A\}$. Binarized scoring rule (BSR) uses two monetary prizes M_h and M_l for payment (where $M_h > M_l \geq 0$), and two i.i.d. draws $X_1, X_2 \sim U[0,1]$ to determine the outcome [Hossain and Okui, 2013, Wilson and Vespa, 2018]. Specifically, if $s = A$ is true, the DM gets the prize M_h so long as their stated belief a is greater than at least one of the two uniform draws X_1 and X_2. If $s = B$ is true, the DM gets the prize M_h so long as their stated belief a is less than at least one of the two uniform draws X_1 and X_2. Otherwise, the DM gets the prize M_l.

Given the true belief p, the probability of winning the better prize M_h is given by

$$\pi(p, a) = p * (1 - (1 - a)^2) + (1 - p) * (1 - a^2) \tag{D.1}$$

Thus, BSR generates a reduced lottery $\mathcal{L}(a|p) = \pi(p, a) \circ M_h \oplus (1 - \pi(p, a)) \circ M_l$. When the true belief p is a singleton and the choice space of a is continuous on $[0, 1]$ (as in the classical method, for example, slider bar), the best response is $a^*(p) = p$ as $\mathcal{L}(a^*(p)|p)$ stochastically dominates any other available lottery $\mathcal{L}(a|p)$.

Consider the situation where the true belief p follows a non-degenerate distribution $f(p)$, with $\mu_p = E(p)$ and $\sigma_p^2 = Var(p) > 0$. The DM can directly select any number between 0 and 1 as in classical methods. The distribution over p captures the idea that the perception of $Pr\{s = A\}$ can be noisy, uncertain, or imprecise [Enke and Graeber, 2023, Frydman and Jin, 2022, Giustinelli et al., 2022]. Without loss of generality, assume $M_l = 0$. Given the true belief p, finding the optimal stated belief $a \in [0, 1]$ that maximizes the expected utility in the BSR is equivalent to maximizing the likelihood of receiving the prize M_h. Unlike the case where the true belief p is a singleton, the optimization problem now involves maximizing the expected likelihood of receiving the prize M_h:

$$\max_a \; E_p[p * (1 - (1 - a)^2) + (1 - p) * (1 - a^2)] \tag{D.2}$$

where a has a continuous choice space between 0 and 1, i.e., $a \in [0, 1]$. The best response in this situation is to select the point $a^*(p)$ where $a^*(p) = E(p) = \mu_p$.

As the DBM allows the DM not only to choose until a single point but also to choose a random variable with a uniform distribution over a range $[a_l, a_u]$, i.e., $a \sim Uniform[a_l, a_u]$, the optimization problem becomes:

$$\max_a \; E_p\{p * E_a[(1 - (1 - a)^2)|p] + (1 - p) * E_a[(1 - a^2)|p]\} \tag{D.3}$$

which is equivalent to

$$\max_a \{-Var(a) - [E(a) - E(p)]^2 + E(1 - p) + [E(p)]^2\} \tag{D.4}$$

where $Var(a)$ and $E(a)$ denote the variance and the mean of stated belief a, respectively. To maximize the expected utility, it is optimal to choose until the point a^* where $a^* = E(p)$ and $Var(a) = 0$.

In sum, given the true belief p, to maximize expected utility, it is optimal to select the point $a^*(p) = E(p)$. This holds true whether the true belief follows a non-degenerate distribution or if the DM is allowed to select a range or a mass point as their belief.

D.2 BDM with Myopic DM

If the DM is myopic, instead of perfectly foreseeing that the optimal choice is to choose until the point a^* where $a^* = E(p)$ and $Var(a) = 0$ in the DBM, they may compare among the three options, followed by an immediate exit in each step: $Uniform[I_l, \frac{I_l+I_u}{2}]$, $Uniform(\frac{I_l+I_u}{2}, I_u]$, or exiting with $Uniform[I_l, I_u]$, where I_l and I_u denote the upper and lower bounds of the interval in each step, respectively. Thus, the likelihood of receiving the prize M_h of choosing each option is:

- when $a = Uniform[I_l, \frac{I_l+I_u}{2}]$:

$$\begin{aligned}
&- Var(a) - [E(a) - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(\frac{I_l+I_u}{2} - I_l)^2}{12} - [\frac{(\frac{I_l+I_u}{2} + I_l)}{2} - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12 * 4} - [\frac{3I_l + I_u}{4} - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12 * 4} - (\frac{3I_l + I_u}{4})^2 + E(p) * \frac{3I_l + I_u}{2} + E(1-p)
\end{aligned} \tag{D.5}$$

- when $a = Uniform(\frac{I_l+I_u}{2}, I_u]$:

$$
\begin{aligned}
&- Var(a) - [E(a) - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - \frac{I_l+I_u}{2})^2}{12} - [\frac{(\frac{I_l+I_u}{2} + I_u)}{2} - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12 * 4} - [\frac{I_l + 3I_u}{4} - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12 * 4} - (\frac{I_l + 3I_u}{4})^2 + E(p) * \frac{I_l + 3I_u}{2} + E(1-p)
\end{aligned}
\tag{D.6}
$$

- when exiting with $a = Uniform[I_l, I_u]$:

$$
\begin{aligned}
&- Var(a) - [E(a) - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12} - [\frac{I_l + I_u}{2} - E(p)]^2 + E(1-p) + [E(p)]^2 \\
&= -\frac{(I_u - I_l)^2}{12} - (\frac{I_l + I_u}{2})^2 + E(p) * (I_l + I_u) + E(1-p)
\end{aligned}
\tag{D.7}
$$

Whenever $E(p) < \frac{I_u+I_l}{2}$, choosing $Uniform[I_l, \frac{I_l+I_u}{2}]$ yields a higher likelihood of receiving M_h than choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$ or exiting with $Uniform[I_l, I_u]$. Similarly, whenever $E(p) > \frac{I_u+I_l}{2}$, choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$ yields a higher likelihood of receiving M_h than choosing $Uniform[I_l, \frac{I_l+I_u}{2}]$ or exiting with $Uniform[I_l, I_u]$. Whenever $E(p) = \frac{I_u+I_l}{2}$, all three options yield the same likelihood of receiving M_h. Thus, the DM would be indifferent in choosing any of the three options.

To sum up, whenever $E(p)$ is strictly within one of the two narrowed intervals, it is optimal to choose the one that contains $E(p)$. Otherwise, the myopic DM is indifferent between choosing $Uniform[I_l, \frac{I_l+I_u}{2}]$, choosing $Uniform(\frac{I_l+I_u}{2}, I_u]$, or exiting with $Uniform[I_l, I_u]$.

Bibliography

Johannes Abeler, Daniele Nosenzo, and Collin Raymond. Preferences for truth-telling. *Econometrica*, 87(4):1115–1153, 2019.

Wifag Adnan, K Peren Arin, Gary Charness, Juan A Lacomba, and Francisco Lagos. Which social categories matter to people: An experiment. *working paper*, 2021.

Marina Agranov and Pietro Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68, 2017.

Marina Agranov and Pietro Ortoleva. Ranges of preferences and randomization. *Report, Princeton Univ.[659]*, 2020.

Marina Agranov and Pietro Ortoleva. Revealed preferences for randomization: An overview. In *AEA Papers and Proceedings*, volume 112, pages 426–430. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2022.

Marina Agranov, Paul J Healy, and Kirby Nielsen. Stable randomisation. *The Economic Journal*, 133(655):2553–2579, 2023.

Ingvild Almås, Alexander W Cappelen, Kjell G Salvanes, Erik Ø Sørensen, and Bertil Tungodden. Fairness and family background. *Politics, Philosophy & Economics*, 16 (2):117–131, 2017.

James Andreoni and B Douglas Bernheim. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636, 2009.

Aurélien Baillon. Eliciting subjective probabilities through exchangeable events: An advantage and a limitation. *Decision Analysis*, 5(2):76–87, 2008.

Michele Belot and Jeroen Van de Ven. How private is private information? the ability to spot deception in an economic game. *Experimental economics*, 20:19–43, 2017.

Daniel J Benjamin, James J Choi, and A Joshua Strickland. Social identity and preferences. *American Economic Review*, 100(4):1913–28, 2010.

Helen Bernhard, Ernst Fehr, and Urs Fischbacher. Group affiliation and altruistic norm enforcement. *American Economic Review*, 96(2):217–221, 2006.

Jean-François Bonnefon, Astrid Hopfensitz, and Wim De Neys. The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, 142(1):143, 2013.

Jean-François Bonnefon, Astrid Hopfensitz, and Wim De Neys. Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*, 26(3): 276–281, 2017.

Jordi Brandts and Gary Charness. The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398, 2011.

Zixing Cai and Yong Wang. A multiobjective optimization-based evolutionary algorithm for constrained optimization. *IEEE Transactions on evolutionary computation*, 10(6): 658–675, 2006.

Alexander W Cappelen, Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden. The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827, 2007.

Alexander W Cappelen, James Konow, Erik Ø Sørensen, and Bertil Tungodden. Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, 103(4):1398–1413, 2013.

Lea Cassar and Arnd H Klein. A matter of perspective: How failure shapes distributive preferences. *Management Science*, 65(11):5050–5064, 2019.

Elena Cettolin and Arno Riedl. Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization. *Journal of Economic Theory*, 181:547–585, 2019.

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

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The quarterly journal of economics*, 117(3):817–869, 2002.

Gary Charness, Luca Rigotti, and Aldo Rustichini. Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352, 2007.

Gary Charness, Ramón Cobo-Reyes, and Natalia Jiménez. An investment game with third-party intervention. *Journal of Economic Behavior & Organization*, 68(1):18–28, 2008.

Gary Charness, Ramon Cobo-Reyes, and Natalia Jimenez. Identities, selection, and contributions in a public-goods game. *Games and Economic Behavior*, 87:322–338, 2014.

Gary Charness, Uri Gneezy, and Vlastimil Rasocha. Experimental methods: Eliciting beliefs. *Journal of Economic Behavior & Organization*, 189:234–256, 2021.

Daniel L Chen, Martin Schonger, and Chris Wickens. otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

Jingnan Chen and Daniel Houser. Promises and lies: can observers detect deception in written messages. *Experimental Economics*, 20:396–419, 2017.

Yan Chen and Sherry Xin Li. Group identity and social preferences. *American Economic Review*, 99(1):431–57, 2009.

Yan Chen, Sherry Xin Li, Tracy Xiao Liu, and Margaret Shih. Which hat to wear? impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, 84:58–86, 2014.

Todd L Cherry, Peter Frykblom, and Jason F Shogren. Hardnose the dictator. *American Economic Review*, 92(4):1218–1221, 2002.

Miruna Cotet and Ian Krajbich. Response times in the wild: ebay sellers take hours longer to reject high offers and accept low offers. *Available at SSRN 3804578*, 2021.

Sergio Currarini and Friederike Mengel. Identity, homophily and in-group bias. *European Economic Review*, 90:40–55, 2016.

Jason Dana, Roberto A Weber, and Jason Xi Kuang. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33: 67–80, 2007.

David Danz, Lise Vesterlund, and Alistair J Wilson. Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 2022.

Bruno Deffains, Romain Espinosa, and Christian Thöni. Political self-serving bias and redistribution. *Journal of Public Economics*, 134:67–74, 2016.

Ambuj Dewan and Nathaniel Neligh. Estimating information cost functions in models of rational inattention. *Journal of Economic Theory*, 187:105011, 2020.

Ruben Durante, Louis Putterman, and Joël Van der Weele. Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, 12(4):1059–1086, 2014.

Nadja Dwenger and Tim Lohse. Do individuals successfully cover up their lies? evidence from a compliance experiment. *Journal of Economic Psychology*, 71:74–87, 2019.

Nadja Dwenger, Dorothea Kübler, and Georg Weizsäcker. Flipping a coin: Evidence from university applications. *Journal of Public Economics*, 167:240–250, 2018.

Catherine C Eckel and Philip J Grossman. Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392, 2005.

Benjamin Enke and Thomas Graeber. Cognitive uncertainty. *The Quarterly Journal of Economics*, 138(4):2021–2067, 2023.

Romain Espinosa, Bruno Deffains, and Christian Thöni. Debiasing preferences over redistribution: An experiment. *Social Choice and Welfare*, 55(4):823–843, 2020.

Armin Falk and Florian Zimmermann. Information processing and commitment. *The Economic Journal*, 128(613):1983–2002, 2018.

Paul Feldman and John Rehbeck. Revealing a preference for mixtures: An experimental study of risk. *Quantitative Economics*, 13(2):761–786, 2022.

Urs Fischbacher and Franziska Föllmi-Heusi. Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547, 2013.

Cary Frydman and Lawrence J Jin. Efficient coding and risky choice. *The Quarterly Journal of Economics*, 137(1):161–213, 2022.

Cary Frydman and Ian Krajbich. Using response times to infer others' private information: An application to information cascades. *Management Science*, 68(4):2970–2986, 2022.

Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–3684, 2018.

Laura K Gee, Marco Migueis, and Sahar Parsa. Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness. *Experimental Economics*, 20(4):894–923, 2017.

Philipp Gerlach, Kinneret Teodorescu, and Ralph Hertwig. The truth about lies: A meta-analysis on dishonest behavior. *Psychological bulletin*, 145(1):1, 2019.

Pamela Giustinelli, Charles F Manski, and Francesca Molinari. Precise or imprecise probabilities? evidence from survey response related to late-onset dementia. *Journal of the European Economic Association*, 20(1):187–221, 2022.

Uri Gneezy. Deception: The role of consequences. *American Economic Review*, 95(1): 384–394, 2005.

Lorenz Goette, David Huffman, and Stephan Meier. The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2):212–216, 2006.

Harrison G Gough. A creative personality scale for the adjective check list. *Journal of personality and social psychology*, 37(8):1398, 1979.

Joshua D Greene and Joseph M Paxton. Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, 106 (30):12506–12511, 2009.

Zachary Grossman. Strategic ignorance and the robustness of social preferences. *Management Science*, 60(11):2659–2665, 2014.

Karla Hoff and Priyanka Pandey. Discrimination, social identity, and durable inequalities. *American economic review*, 96(2):206–211, 2006.

Tanjim Hossain and Ryo Okui. The binarized scoring rule. *Review of Economic Studies*, 80(3):984–1001, 2013.

Tiffany A Ito, Jeff T Larsen, N Kyle Smith, and John T Cacioppo. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of personality and social psychology*, 75(4):887, 1998.

Ting Jiang. Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, 93:328–336, 2013.

Daniel Kahneman. *Thinking, fast and slow*. macmillan, 2011.

Arkady Konovalov and Ian Krajbich. Decision times reveal private information in strategic settings: Evidence from bargaining experiments. *The Economic Journal*, 133(656): 3007–3033, 2023.

James Konow. Fair shares: Accountability and cognitive dissonance in allocation decisions. *American economic review*, 90(4):1072–1091, 2000.

Kai A Konrad, Tim Lohse, and Salmai Qari. Deception choice and self-selection–the importance of being earnest. *Journal of Economic Behavior & Organization*, 107: 25–39, 2014.

Jonathan Lafky, Ernest K Lai, and Wooyoung Lim. Preferences vs. strategic thinking: An investigation of the causes of overcommunication. *Games and Economic Behavior*, 136:92–116, 2022.

Lars J Lefgren, David P Sims, and Olga B Stoddard. Effort, luck, and voting for redistribution. *Journal of Public Economics*, 143:89–97, 2016.

Geoffrey J Leonardelli and Marilynn B Brewer. Minority and majority discrimination: When and why. *Journal of Experimental Social Psychology*, 37(6):468–485, 2001.

Yucheng Liang. Learning from unknown information sources. *Available at SSRN 3314789*, 2022.

Charles F Manski. Measuring expectations. *Econometrica*, 72(5):1329–1376, 2004.

Kendra N McLeish and Robert J Oxoby. Identity, cooperation, and punishment. 2007.

Bin Miao and Songfa Zhong. Probabilistic social preference: how machina's mom randomizes her choice. *Economic Theory*, 65:1–24, 2018.

Kirby Nielsen and Luca Rigotti. Revealed incomplete preferences. *Available at SSRN 4622145*, 2023.

Robert J Oxoby and John Spraggon. Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization*, 65(3-4):703–713, 2008.

Rong Rong, Daniel Houser, and Anovia Yifan Dai. Money or friends: Social identity and deception in networks. *European Economic Review*, 90:56–66, 2016.

Ariel Rubinstein. Irrational diversification in multiple decision problems. *European Economic Review*, 46(8):1369–1378, 2002.

Ariel Rubinstein. A typology of players: Between instinctive and contemplative. *The Quarterly Journal of Economics*, 131(2):859–890, 2016.

Marta Serra-Garcia and Uri Gneezy. Nonreplicable publications are cited more than replicable ones. *Science advances*, 7(21):eabd1705, 2021.

Shaul Shalvi, Ori Eldar, and Yoella Bereby-Meyer. Honesty requires time (and lack of justifications). *Psychological science*, 23(10):1264–1270, 2012.

Moses Shayo. Social identity and economic policy. *Annual Review of Economics*, 12: 355–389, 2020.

Martin Storme, Pinar Celik, and Nils Myszkowski. Creativity and unethicality: A systematic review and meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 15(4):664, 2021.

Gustav Tinghög, David Andersson, and Daniel Västfjäll. Are individuals luck egalitarians?–an experiment on the influence of brute and option luck on social preferences. *Frontiers in psychology*, 8:460, 2017.

Alexander Todorov, Christopher Y Olivola, Ron Dotsch, and Peter Mende-Siedlecki. Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual review of psychology*, 66:519–545, 2015.

Philippe PFM Van de Calseyde, Gideon Keren, and Marcel Zeelenberg. Decision time as information in judgment and choice. *Organizational Behavior and Human Decision Processes*, 125(2):113–122, 2014.

Alistair Wilson and Emanuel Vespa. Paired-uniform scoring: Implementing a binarized scoring rule with non-mathematical language. Technical report, Working paper, 2018.

Matthew Wiswall and Basit Zafar. How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169, 2015a.

Matthew Wiswall and Basit Zafar. Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824, 2015b.

Michael Woodford. Stochastic choice: An optimizing neuroeconomic model. *American Economic Review*, 104(5):495–500, 2014.