Essays on Behavioral and Experimental Economics

Yaoyao Xu

School of Economics

University of Edinburgh

This dissertation is submitted for the degree of

Doctor of Philosophy

2023

Abstract

In this dissertation of three chapters, I study individuals' strategic sophistication in decision-making, specifically level-k reasoning and forward-looking behavior. The first chapter studies subjects' iterative reasoning ability to derive equilibrium and effectiveness of teaching in improving their decisions (if subjects have bounded reasoning ability). In the experiment, I first measure subjects' iterative reasoning ability using rational computer opponents to control strategic uncertainty from beliefs about opponents' moves. I find that 90% of the subjects cannot play equilibrium. Next, I randomly treat subjects with a tutorial about iterative reasoning, and I find that 45% of the subjects learned to play equilibrium. Finally, I study subjects' perceived ability and improvement by asking subjects to assess their performance. I find that low-ability subjects who do not learn have persistent misperception (i.e., overestimation) of their own performance and learning ability to improve performance. The findings in this chapter suggest that training on strategic thinking is a cost-effective intervention to improve decisions and to aid policy implementation, and low-ability subjects need a more intensive treatment.

The first chapter shows that subjects have limited reasoning ability, which might explain the well-documented bounded rationality in games. However, it is still unclear at the individual level whether the low-level of rationality is due to limited reasoning ability or low-order beliefs about opponents' rationality. The second chapter reports a within-subject experiment on Amazon Mechanical Turk (Mturk), where subjects played ring games against two types of opponents simultaneously, other Mturk subjects and themselves. Lk players who are bounded by their ability would display the same reasoning depth when facing either type of the opponent (ability-bounded Lk), otherwise the players would perform higher reasoning depth when playing against themselves than other participants (belief-bounded Lk). I find that 76% of them are ability-bounded Lk

players while only 10% are belief-bounded Lk players, indicating that limited reasoning ability is the primary determinant of bounded rationality.

The third chapter provides an experimental investigation of the evolutionary game model (Oyama et al., 2015) which predicts transitions among strict Nash equilibria under inexact (inaccurate but unbiased) information of opponents' behaviors. We design a quasi-continuous-time experiment with two treatments differing in information accuracy. A group of subjects played a coordination game repeatedly in either treatment. We observe more efficiency-improving transitions among strict Nash equilibria in the more accurate information treatment than in the less accurate information treatment, contrary to the theory. We further find that more accurate information about opponents' behaviors induces more subjects to engage in forward-looking behavior, i.e., persistent strategic deviations from the myopic best responses to the information received, which facilitates efficiency-improving equilibrium transitions. When information is less accurate, subjects are less responsive to changes in the information. The slow response to the information either blocks or delays efficiency-improving equilibrium transitions.

Lay Summary

This dissertation consists of three chapters, studying how people make decisions in strategic interactions using controlled lab and online experiments.

When people interact with others, their decisions potentially depend on both their own reasoning ability and their beliefs about the decisions of those they interact with. When people are making non-optimal decisions, is it due to people's limited ability or their responses to beliefs that others are making non-optimal decisions? The first chapter investigates questions concerning people's reasoning ability to identify optimal strategies in games and the effectiveness of teaching in improving their ability. In this experiment, subjects played a set of ring games (introduced by Kneeland, 2015) against highly cognitively sophisticated computer players. This design allows me to measure subjects' iterative reasoning ability without strategy uncertainty from subjects' beliefs that their opponents are making mistakes. I find that most of my subjects have bounded reasoning ability. Next, I randomly provide my subjects with a tutorial on how to derive the optimal strategies. I find that 45% of the subjects learn to identify the optimal strategies after the tutorial. These findings in the chapter are not just important for improving individual strategic decision-making, but also important for policy implementation, since the results from this chapter suggest that to achieve the intended outcome of new policies such as taxes, most people need guidance to react to the policies in the desired way. While some individuals can learn to calculate the optimal strategies, others may need more help than others.

While the first chapter suggests that limited ability could be a reason for non-optimal decisions, it cannot determine the primary driver behind those decisions at the individual level. The second chapter investigates this question by conducting an online experiment on Amazon Mechanical Turk (Mturk). In the experiment, subjects played a set of ring games

against two types of opponents, themselves and other Mturk subjects simultaneously. When playing against others, subjects need to form beliefs about others' actions and have the reasoning ability. When playing against themselves, it does not involve belief formation and thus only requires the ability. Therefore, if subjects' decisions are driven by ability, they would make the same decisions when facing either opponent. Instead, if their decisions are driven by belief, they would behave differently. I find 76% of my subjects' decisions can be explained by limited ability while 10% can be explained by belief. The result indicates that subjects' choices are primarily bounded by their reasoning ability rather than beliefs.

The third chapter provides an experimental investigation of a theory (Oyama et al., 2015), which suggests that subjects can coordinate to transit from one social convention (modelled as an equilibrium) to another more efficient social convention (another equilibrium) when they receive signals with some information about others' behaviors. In the experiment, a group of subjects played a coordination game repeatedly in one of the two treatments differing in the amount of information about others' behaviors. Different from the theory, we find that subjects transit to the most efficient equilibrium more often with more information. The transitions are facilitated by subjects who intentionally deviate from the best responses to the signals, and play an action favoring a transition to a more efficient equilibrium. More information induces more such subjects, resulting in more transitions to the most efficient equilibrium. Meanwhile, with less information, subjects are less responsive to the change in the signals. Such behavior blocks or delays the transitions to a more efficient equilibrium.

Acknowledgements

I thank my supervisors Tatiana Kornienko and Ed Hopkins. This dissertation is possible thanks to their support, guidance and encouragement in the past five years.

I also thank Mariann Ollar, Kohei Kawamura, Eugenio Proto, Andrew Clausen, David Comerford, Larbi Alaoui, Jonna Olsson, Ina Taneva, Stuart Breslin, and the audience at various seminars and conferences for the valuable discussion and comments. I thank my coauthors Zhi Li and Jianxuan Lyu for their contributions to Chapter 3 of this dissertation. Finally, I thank Alex Kostylev for his assistance to run the experiments and for emotional support.

I acknowledge financial support from the School of Economics and Moray Endowment Fund at the University of Edinburgh, and the Scottish Economic Society, and acknowledge that ChatGPT was used for grammar checking.

Contents

List of Figures

List of Tables

Chapter 1

Can Iterative Reasoning be Taught in the Lab?

1.1 Introduction

The two fundamental assumptions often made in game theory are rationality and common knowledge of rationality. Deriving equilibrium of a game usually requires people to think iteratively through a process "you think that I think that you are rational...". Depending on the game, the number of steps of such iterative reasoning required to calculate the equilibrium can be indefinite, and people's initial responses often turn out to be non-optimal or non-equilibrium (see Crawford et al., 2013, for a review). Non-optimal responses could cause people to leave money on the table in various strategic situations, such as negotiations and auctions. It could further cause difficulty in the delivery of the intended outcomes when a new policy or institutional design is introduced, such as tax. To improve decision-making and aid policy implementation, it is critical to understand people's ability to do iterative reasoning, and to have a cost-effective way to improve ability.

By focusing on reasoning ability, this paper differs from much of the existing literature which focuses on explaining non-equilibrium play (or bounded rationality) using non-equilibrium beliefs (e.g., Slonim, 2005; Palacios-Huerta and Volij, 2009; Agranov et al., 2012; Georganas et al., 2015; Gill and Prowse, 2016), as in level-k model (Stahl and

Wilson, 1994, 1995; Nagel, 1995; Ho et al., 1998). The model starts with $L0$ players, who are assumed to be irrational and play randomly or according to a simple rule. Lk players best respond to belief that all the opponents are $L(k-1)$ players, for any $k \geq 1$. That is, the model implicitly assumes that players differing in reasoning depths are due to beliefs that opponents are one level below themselves, instead of their own limitation in intrinsic ability of reasoning.[1]

Thus, this paper takes a step back and studies to what extent subjects' reasoning ability is bounded, and further whether subjects' ability can be improved through teaching to play equilibrium. I also look at whether subjects can accurately assess their ability and improvement in ability by comparing their actual and self-assessed performance before and after the teaching.

I investigate these questions in a lab experiment using simultaneous 4-player dominance solvable ring games (introduced by Kneeland, 2015). Subjects played games against three rational computer players before and after a tutorial (without any feedback), followed by assessment of their own performance in games.

It is possible that subjects believe their opponents have much lower ability, and so decide to best respond to their beliefs. As a result, the reasoning depth they display is lower than their true capability. To control for the strategic uncertainty stemming from such belief, the rational computer opponents are introduced. Subjects were informed that the computer players are rational and always play best responses. Thus subjects were assured that their opponents are highly sophisticated, and provided incentive to reason until reach their capability. In effect, this turns the four-player game into one-person decision problem. More importantly, this design allows one to identify the whole distribution of iterative reasoning ability (Ak), up to four steps of iterative reasoning ($A4$).

[1] Cognitive hierarchy model introduced by Camerer et al. (2004) is another prominent work which assumes Lk player best respond to a distribution of all the lower types.

To study whether subjects can improve their reasoning ability from teaching, I design two novel tutorials. The tutorial in the treatment group teaches iterative dominance, *relevant* to solve the games, while the one in the control group teaches color subtraction and addition, *irrelevant* to strategic decision making, or even general logical thinking. Thus, the irrelevant tutorial serves to control for the earlier exposure to the games and a possibility of introspective learning. It has been shown that even without any feedback subjects could still learn from experience or introspection (e.g., Weber, 2003; Puente, 2020, among others). So it is essential to have a control group to avoid potential confounded effect of the relevant tutorial.

Finally, to investigate subjects' assessment accuracy of their reasoning ability and ability improvement, subjects in the control and treatment groups were asked to report their perceived performance in games both before and after the tutorials. It is reasonable to expect that subjects would have a judgement that their post-tutorial performance is better than pre-tutorial performance, mostly when teaching *and/or* introspection made them understand something they believe useful in their improving performance. Further, it is reasonable to assume that only if the treatment group reports greater perceived performance improvement than the control group, the perceived improvement can be attributed to learning from teaching in addition to experience. That is, similar to examining actual performance improvement, it is of importance to account for the perceived improvement from experience.

The first contribution of this paper is to provide an empirical distribution of ability bounds. I find that before the tutorials, 44% of our subjects are $A0$, 34% are $A1$, 9% are $A2$, 3% are A3, and only 10% played equilibrium and thus are $A4$.

Thus, this paper documents the extent of limited reasoning ability as a determinant of level-k behavior. By using rational computer players, the belief formation is removed and the ability being measured is the reasoning depth that subjects can perform to calculate the best responses iteratively. This ability definition is different from the most papers in the literature on this topic. For example, Bosch-Rosa and Meissner (2020) look at whether or not subjects have the ability to solve one-player guessing game, which is effectively a math question, but they do not look at subjects' ability measures. Alaoui and Penta (2016); Alaoui et al. (2020); Jin (2021) look at the maximum *level* a subject can conceive of, which consists of the ability of forming higher-order beliefs and calculating iterated best responses. Friedenberg et al. (2021) employ a more general definition of ability of iterative reasoning, which is not restricted to best responses.[2]

Puente (2020) explores similar idea of ability to this paper, and finds 74% $A4$ subjects, a substantially higher share of high-ability subjects, compared to only 10% in my sample. However, his experiment has two features which could lead to overestimation of ability bounds. First, his subjects played the ring games against rational computer players sequentially, instead of simultaneously. It is possible that the sequential design imposes procedural rationality (Simon, 1955, 1976, 1982), i.e., inform subjects of correct reasoning path. In contrast, the simultaneous design used in this paper would require subjects to identify the reasoning path themselves. Second, the payoff tables in Puente (2020) were adopted from Kneeland (2015), which have best responses partially coincide with risk dominant actions, potentially leading to misidentify low types into high types (Jin, 2022; Sprenger and Zhao, 2021). Therefore, in order to accommodate the problem of misidentification, in my experiment I adopted the payoff matrices of Jin (2021) (before the tutorials) and developed new matrices (after the tutorials), in which risk dominant actions and best responses are separated. Because of high proportion of subjects playing equilibrium in Puente (2020)'s experiment, he concludes that reasoning ability has a very limited role in bounded rationality. Nevertheless, my findings on reasoning ability shed

[2] They study the ability to go through the iterative process that "you think that I think that you think ..", where the response can be either rational or not. It differs from the process considered in this paper and other papers in the literature that "you think that I think that you are *rational*" and respond *rationally*.

new lights on explanation of non-equilibrium behaviour, and highlights the importance of bounded reasoning ability.

The second contribution of this paper is to provide a cost-effective intervention to improve reasoning ability via the relevant tutorial, and a simple control for introspective learning from experience by using the irrelevant tutorial. I find that indeed some subjects in the control group make ability improvement, though on average experience only has an insignificant effect. This emphasizes the importance and necessity of the control group. After experience is accounted for, I find that the relevant tutorial has a significantly positive effect on reasoning ability. In particular, 45% of the subjects in the treatment group learn to play equilibrium. More importantly, I find that high-cognition subjects not only have higher reasoning ability prior to tutorials, but also tend to make larger ability improvement after the tutorials.

These findings fit into a broad literature on teaching economics (see Allgood et al., 2015, for a review), and is most relevant to the works which study whether subjects make better strategic decisions after training or teaching (e.g., Johnson et al., 2002; James and Cohen, 2004; Martins et al., 2017; Verbrugge et al., 2018; Fan, 2000; Costa-Gomes and Crawford, 2006). In particular, Alaoui et al. (2020) also teach subjects to do iterative reasoning in 11-20 games. They find that about half of their subjects switch to equilibrium after the teaching. Different from this paper, their subjects played games against human opponents. As a result, they do not quantify improvement in ability to calculate iterated best responses.

The third contribution of this paper is to study how subjects assess or perceive performance and performance improvement. To the best of my knowledge, this paper is the first to study perceived improvement experimentally by asking subjects to predict their absolute performance before and after the tutorials. After controlling for perceived improvement via experience, I find that subjects perceive their performance improved after the relevant tutorial even when they, in fact, have not improved. It suggests a possibility of overconfidence in learning ability. Importantly, this result can also be interpreted as perceived teaching effectiveness in improving ability, as subjects would believe that they improve only when the relevant tutorial (in addition to experience) made them understand something useful. Thus the results suggest that even subjects who do not improve think the relevant tutorial is effective in improving their performance.[3] More generally, the results on teaching iterative reasoning and perceived teaching effectiveness also demonstrate the value of education in improving decisions, and people's general perception of education usefulness. Importantly, education of strategic thinking might establish larger betterment in the long run if it can be incorporated into pre-universities education.

Furthermore, I find that both low- and high-ability subjects tend to make inaccurate assessments of their own performance but in the opposite direction: low-ability subjects tend to overestimate while high-ability subjects tend to underestimate, the observation known as Dunning-Kruger effect (Kruger and Dunning, 1999). Such tendency cannot be corrected by the exposure to the relevant tutorial if subjects do not learn to play equilibrium. That is, low-ability subjects who do not learn from the treatment exhibit *multi-dimensional overconfidence* in both performance and learning ability to improve performance.

These findings relate the literature on self-assessment (e.g., Wason, 1968; Dominitz, 1998; Kruger and Dunning, 1999; Dominitz, 1998; Fischhoff et al., 2000; Grimes, 2002; Nowell and Alston, 2007; Schlösser et al., 2013; Grossman and Owens, 2012; Thoma, 2016; also see Dunning et al., 2004; Andrade, 2019, for reviews), and effect of teaching

[3]The perceived teaching effectiveness may contribute to the literature on Student Evaluation of Teaching (SET) (see Spooren et al., 2013; Allgood et al., 2015, for reviews), which is the predominant measure of perceived teaching effectiveness used by most colleges and universities. The perception of course instructors' characteristics such as enthusiasm and presentation skill (Siegfried and Walstad, 1998) are found to affect SET scores. The measure used in this paper is to a less subjective extent than SET in the sense that these subjective factors are controlled in a lab experiment.

on accuracy of self-assessment (e.g., Ferraro, 2010; Hossain and Tsigaris, 2015; Guest and Riegler, 2017). The main contribution of this paper to the above literature is to study perceived performance and improvement in the strategic environment, specifically under level-k framework. In particular, I document two types of overestimation or overconfidence, and thus also contribute to the literature on overconfidence (e.g., see Moore and Healy, 2008; Moore and Schatz, 2017, for reviews).

The remainder of the paper is structured as follows. Section 1.2 describes ring game and reasoning depth identification strategy. The experimental design is presented in section 1.3. Hypotheses are summarized in section 1.4. Section 1.5 and 1.6 report results. Finally, section 1.7 concludes.

1.2 Ring Game and Identification Strategy

Ring game was first introduced by Kneeland (2015) to study rationality bounds Lk, defined by the maximum reasoning steps a subject would perform when playing games against other human subjects. It is a 4-player simultaneous game with three actions and the following structure: the payoff of Player 1 depends only on the decisions of Player 1 and Player 2; the payoff of Player 2 depends on the decisions of Player 2 and Player 3; the payoff of Player 3 depends on the decisions of Player 3 and Player 4; the payoff of Player 4 depends on the decisions of Player 4 and Player 1 (see e.g., G1 and G2 of Figure 1.1 taken from Jin (2021)). Importantly, Player 4 has a dominant strategy which determines the best responses of the other three players iteratively, and so the game is dominance solvable.

Another important feature of ring games is that there is a pair of 4-player ring games, where Player 1, 2 and 3 have the same payoff tables while Player 4 does not. Specifically, the rows of Player 4's payoff matrices in the two 4-player games are swapped so that the dominant strategy has a different label, leading to different iterated best responses of the

Player 1's Payoffs (Player 2's actions)

Player 1's actions	d	e	f	
a	30	6	20	(56)
b	4	20	32	(56)
c	12	40	6	(58)

Player 2's Payoffs (Player 3's actions)

Player 2's actions	g	h	i	
d	32	24	4	(60)
e	36	20	10	(66)
f	26	18	16	(60)

Player 3's Payoffs (Player 4's actions)

Player 3's actions	j	k	l	
g	18	20	26	(64)
h	14	8	36	(58)
i	6	24	28	(58)

Player 4's Payoffs (Player 1's actions)

Player 4's actions	a	b	c	
j	8	12	24	(44)
k	20	16	32	(68)
l	18	14	20	(52)

G1

Player 1's Payoffs (Player 2's actions)

Player 1's actions	d	e	f	
a	30	6	20	(56)
b	4	20	32	(56)
c	12	40	6	(58)

Player 2's Payoffs (Player 3's actions)

Player 2's actions	g	h	i	
d	32	24	4	(60)
e	36	20	10	(66)
f	26	18	16	(60)

Player 3's Payoffs (Player 4's actions)

Player 3's actions	j	k	l	
g	18	20	26	(64)
h	14	8	36	(58)
i	6	24	28	(58)

Player 4's Payoffs (Player 1's actions)

Player 4's actions	a	b	c	
j	18	14	20	(52)
k	8	12	24	(44)
l	20	16	32	(68)

G2

Notes: The best responses are in red. Payoff sums for each action are in the parentheses with the highest sums (indicating risk dominance) underlined.

Figure. 1.1: The Ring Games I (Jin, 2021): Pre-tutorial.

other three players in the two games. Thus, the probability that a subject chooses best responses at a particular player's position in both games by chance is relative low.

This feature is essential for Kneeland (2015)'s Natural Exclusive Restriction for the empirical identification of subjects' types: because of the same payoff tables in two games, a Lk subject is assumed to choose *the same action* in both G1 and G2 at the positions which require higher reasoning steps than k for $k \geq 1$.[4] For example, a $L1$ player would choose the dominant strategies as Player 4 in the paired games, but would choose the same action in both games as Player 3 due to the same payoff tables i.e., either g, h or i in both G1 and G2. Similarly, (s)he would choose either d, e or f as Player 2 and choose either a, b or c as Player 1 in the two games. Likewise, a $L2$ player would figure out the

[4] The restriction assumes that low types do not respond to the change of high types payoff information. See Kneeland (2015) for a detailed discussion of this assumption.

Table 1.1: Predicted Action Profiles.

Reasoning Depth		Player 1	Player 2	Player 3	Player 4
Games I	4 Steps	(b, a)	(f, d)	(i, h)	(k, l)
	3 Steps	(a, a)/(b, b)/(c, c)	(f, d)	(i, h)	(k, l)
	2 Steps	(a, a)/(b, b)/(c, c)	(d, d)/(e, e)/(f, f)	(i, h)	(k, l)
	1 Step	(a, a)/(b, b)/(c, c)	(d, d)/(e, e)/(f, f)	(g, g)/(h, h)/(i, i)	(k, l)
Games II	4 Steps	(c, a)	(f, e)	(i, g)	(j, l)
	3 Steps	(a, a)/(b, b)/(c, c)	(f, e)	(i, g)	(j, l)
	2 Steps	(a, a)/(b, b)/(c, c)	(d, d)/(e, e)/(f, f)	(i, g)	(j, l)
	1 Step	(a, a)/(b, b)/(c, c)	(d, d)/(e, e)/(f, f)	(g, g)/(h, h)/(i, i)	(j, l)

Notes: The first entry is the action in G1 while the second entry is the action in G2.

(iterative) best responses as Player 3-4 which differ in the two games, but choose the same action in both games as Player 1 or 2. Under the Natural Exclusive Restriction, $L2$ subjects thus can be separated from $L1$ subjects due to different predicted actions at Player 3's position. The identification of $L3$ and $L4$ subjects can be done in the same way: a $L3$ player would choose (iterated) best responses as Player 2-4 but choose either a, b or c as Player 1 in both games. Finally, a $L4$ (or above) player would play equilibrium. The predicted action profile of each type is presented in the top panel of Table 1.1.

In this paper, the ability bound Ak of a subject is defined by the maximum steps of iterative reasoning that this subject would perform when playing games against known rational computer players. Identification of ability bounds Ak closely follows Kneeland (2015)'s assumption and strategy described above, and thus gives the same predicted profile in the pre-tutorial games (shown in Figure 1.1) in the top panel of Table 1.1.

Importantly, I adopt payoff matrices by Jin (2021) because in Kneeland (2015), some best responses coincide with risk dominance action (i.e., the action has the highest payoff sum across columns) and suboptimal actions contain zero payoff. It has been shown that the above two features could lead to misidentification problem (Jin, 2022; Sprenger and Zhao, 2021). Jin (2021)'s matrices remove these two features. Building on that, I

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f	
a	20	30	6	(56)
b	26	12	20	(58)
c	20	4	32	(56)

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i	
d	20	34	10	(64)
e	24	32	4	(60)
f	18	26	16	(60)

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l	
g	8	10	36	(54)
h	20	12	28	(60)
i	24	4	26	(54)

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c	
j	20	32	16	(68)
k	18	24	14	(56)
l	8	22	12	(42)

G1

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f	
a	20	30	6	(56)
b	26	12	20	(58)
c	20	4	32	(56)

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i	
d	20	34	10	(64)
e	24	32	4	(60)
f	18	26	16	(60)

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l	
g	8	10	36	(54)
h	20	12	28	(60)
i	24	4	26	(54)

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c	
j	8	22	12	(42)
k	18	24	14	(56)
l	20	32	16	(68)

G2

Notes: The best responses are in red. Payoff sum of each action is in the parenthesis with the highest sum (indicating risk dominance) underlined.

Figure. 1.2: The Ring Games II: Post-tutorial.

develop an additional set of payoff tables (Figure 1.2), as post-tutorial games, in which the predicted action profiles are presented in the bottom panel of Table 1.1.

1.3 Experimental Design

The experiment consists of a control and a treatment group, each of which has 5 parts. In both groups, Part 1 is designed to provide the baseline measures of subjects' iterative reasoning ability. In this part, subjects played the paired ring games shown in Figure 1.1 in each player position for a total of 8 rounds. Subjects were informed that in any position, their opponents are 3 rational computer players, who are programmed to maximize their own payoff (thus always play iterated best responses) and believe that the other 3 players including the human player are rational payoff maximizers. Thus the strategic uncertainty about opponents' rationality is minimized, allowing for the identification of subjects' ability bounds. Further, both the subjects' roles and the order of the payoff tables were random in each round, and subjects did not receive any feedback on the historical play until the end of the experiment.

In Part 2, subjects faced the most innovative feature of the experiment, which involves a tutorial. The *relevant* tutorial in the treatment group was designed to teach subjects how to do iterative dominance by explaining how 4 payoff-maximising computer players would play the games in Part 1, and providing the equilibrium strategies to the games. The *irrelevant* tutorial in the control group instead taught subjects colour addition and subtraction which provided no useful principles to solve the ring games, so it is expected to have no effect on the reasoning ability. In addition, the irrelevant tutorial was designed to involve sufficiently comprehensive information, similar amount of text, figures and layout to the relevant tutorial. Thus, the irrelevant tutorial ensured that subjects in both control and treatment groups experienced similar duration of the experiment (and thus have similar monetary incentives). More importantly, the color tutorial enables one to control introspective learning from experience due to its irrelevance to the games.[5] Finally, subjects were asked to spend at least 5 minutes and at most 10 minutes on the tutorials.[6]

Part 3 was the same as Part 1 except it involved different paired ring games (see Figure 1.2) which were developed to have the same features of the games as in Jin (2021). This part allows one to measure subjects' performance after the tutorials and so test the effect of the tutorials on the reasoning ability.

In Part 4, subjects were asked to report the number of rounds ($\hat{x}$) they have chosen the best responses given the actions chosen by their opponents in Part 1 and Part 3 in the payoff maximizing manner. This part allows one to test the perceived performance

[5]See Appendix A.4 and A.5 for details of the tutorials.

[6]The maximal time of 10 minutes is chosen based on the pilot session where almost all the subjects managed to finish this part within 10 minutes.

and improved performance via the relevant tutorial. Finally, in Part 5 subjects answered a survey including demographics, and cognitive reflection test (Toplak et al., 2011).

The payoffs in the experiment were expressed in terms of tokens with the exchange rate of £0.15/token. For each of Part 1 and Part 3, one out of eight rounds was randomly selected for payment for playing games. Engagement in tutorials in Part 2 was not paid to motivate learning. Each assessment in Part 4 was incentivized according to the quadratic scoring rule $\max\{0, [3 - 0.3(x - \hat{x})^2]\}$ where x is the actual number of rounds. Furthermore, a flat rate of £3 was paid for answering the survey in Part 5. Finally, subjects were paid £5 for showing up.

1.4 Hypotheses

In the following Ak_{ijg} denotes the ability bound of subject i, and x_{ijg} and $\hat{x}_{ijg}$ denotes the actual and self-assessed number of rounds the subject chooses best responses in Part $j \in \{1, 3\}$ of group $g \in \{C, T\}$ (C for control and T for treatment) respectively.

The first hypothesis concerns iterative reasoning ability assumption in level-k literature as discussed in section 1.1. Testing the hypothesis is straightforward - subjects having ability systematically lower than $A4$ would lead to the rejection of the hypothesis of subjects' sophisticated reasoning ability. Note that in the setting of 4-player ring games, where solving the games requires the maximum reasoning steps of 4, so one cannot distinguish subjects who are only able to do 4 steps from subjects who have higher ability.

Hypothesis 1.1. *Subjects can do at least 4 steps of iterative reasoning when playing against rational computer players: all the subjects will be identified as* $A4$*.*

The second hypothesis assumes subjects could learn from earlier exposure to games. A positive learning effect from experience will be evident if subjects' ability increases after the irrelevant tutorial in the control group.

Hypothesis 1.2. *The experience of playing games in Part 1 has a positive effect on reasoning ability in Part 3 on average for the control group with the irrelevant tutorial:* $Ak_{1C} > Ak_{3C}$.

To test if the relevant tutorial has a positive effect or not, one needs to account for potential introspective learning from experience, and to test if there is larger ability improvement in the treatment group compared to the control group on average.

Hypothesis 1.3. *The relevant tutorial in the treatment group has a positive effect on reasoning ability on average:* $Ak_{3T} - Ak_{1T} > Ak_{3C} - Ak_{1C}$.

The last two hypotheses look at subjects' assessment of performance and improvement. First I hypothesize that subjects tend to overestimate their performance (although sometimes they underestimate), consistent with the literature on overconfidence (e.g., see Moore and Healy, 2008; Moore and Schatz, 2017, for reviews):

Hypothesis 1.4. *Subjects overestimate their performance:* $\hat{x} > x$.

Finally, I assume that regardless of actual ability improvement, subjects tend to believe they have improved, in which case, subjects would report a larger number of rounds that they had made optimal choices in Part 3 than Part 1, i.e., $\hat{x}_3 > \hat{x}_1$. When subjects in the control group do so, i.e., $\hat{x}_{3C} > \hat{x}_{1C}$, this would be due to experience. However, when subjects in the treatment group do so, i.e., $\hat{x}_{3T} > \hat{x}_{1T}$, this could be due to experience and/or the relevant tutorial. To see if the relevant tutorial treatment plays a role, one needs to verify whether the treatment group believes that they have improved more than the control group, similar to the test of the effect of relevant tutorial on reasoning ability (i.e., difference-in-difference approach). Importantly, this hypothesis concerns subjects' confidence in *ability to learn*, a different dimension of confidence in absolute level of performance, mentioned in Hypothesis 1.4. An observation of subjects reporting improvement without actual ability improvement could be an indication of overconfidence in ability to learn.

Hypothesis 1.5. *Regardless of actual ability improvement, subjects in the treatment group believe they have performance improvement resulting from the relevant tutorial:* $\hat{x}_{3T} - \hat{x}_{1T} > \hat{x}_{3C} - \hat{x}_{1C}$ *for both* $Ak_{3T} > Ak_{1T}$ *and* $Ak_{3T} \leq Ak_{1T}$.

1.5 Experimental Results: Learning

The experiment was conducted in person at the Behavioural Laboratory at the University of Edinburgh (BLUE) in 2019. A total of 157 subjects were recruited through BLUE Research Participation System. 80 subjects were in the treatment group while 77 subjects were in the control group. Each subject only participated in one session. Instructions were read aloud in front of the subjects, and a comprehension quiz was carried out to ensure the subjects understand the instructions. To start the game subjects had to answer all quiz questions correctly. While taking the quiz, subjects were allowed to refer to paper instructions and resubmit their answers if they made mistakes. Subjects were not allowed to interact with each other or take notes during the session. To assist them in making decisions, the computer interface allowed subjects to click every cell of each payoff table, and highlight the relevant columns and rows.

The sessions lasted for one hour on average, and the average payment was £14.63 including the show-up fee of £5. The treatment and control group turned out to be balanced in terms of demographics including gender, majoring in economics, study level, and age. See Table A.3 for summary statistics of the sample.

1.5.1 Ability Bounds before the Tutorials

In this section, subjects are classified into types based on their action choices in Part 1 before the tutorials. Following the identification strategy discussed in section 1.2, a subject is classified as type Ak if his or her action profile is exactly matched the predicted

Table 1.2: Distribution of ability bounds in Part 1 before the tutorials.

		$A0$	$A1$	$A2$	$A3$	$A4$	Total
Control	Count	34	25	7	2	9	77
	Percent	44.16	32.47	9.09	2.60	11.69	100.00
Treatment	Count	35	28	7	3	7	80
	Percent	43.75	35.00	8.75	3.75	8.75	100.00
Total	Count	69	53	14	5	16	157
	Percent	43.95	33.76	8.92	3.18	10.19	100.00

action profile of that type. Following Kneeland (2015) I allow for one deviation from the predicted action profile, and if subjects can be classified into two types with one error, subjects are assigned to the lower type conservatively.[7] If subjects made more than one deviation from the predicted profiles, they are assigned to $A0$.[8]

As reported in Table 1.2 in total about 44% of the subjects are $A0$, 34% are $A1$, 9% are $A2$, 4% are $A3$, 10% are $A4$. Such a large proportion of subjects who cannot do 4 steps rejects the hypothesis that subjects are strategically sophisticated in terms of own reasoning ability on average ($Ak < 4$, t-test $p < 0.01$). Note that there is no significant difference in types between the control group and the treatment group (Mann-Whitney test $p = 0.8914$).

Result 1.1. *About 90% of the subjects cannot solve the games and are ability bounded. The remaining 10% played the equilibria and can iteratively reason at least 4 steps.*

[7]Table A.4 shows type classification if subjects are assigned to the higher type when they can be identified as two types with an error. All the results presented below do not change qualitatively if this classification rule is used.

[8]Table A.5 reports the type classification allowing for two errors. The one-error and two-error classifications rules give qualitatively the same results.

1.5.2 Learning from Experience

Before studying the effect of learning from the relevant tutorial on performance in Part 3, let us first look at the learning from experience.

The left panel of Table 1.3 illustrates the change in reasoning ability from Part 1 (before the tutorials) to Part 3 (after the tutorials) at the individual level in the control group which was exposed to the irrelevant color tutorial. 56% of subjects are classified as the same types before and after the color tutorial. Wilcoxon matched-pairs signed-rank test suggests that the difference between the two parts is only marginally significant at 10% level ($p = 0.057$), suggesting experience does not have a significant impact on reasoning ability on average.

At the individual level, 29% of the subjects switched to higher types in Part 3. Note that two subjects who are identified as $A0$ in Part 1, jump to $A4$ in Part 3. This indicates that there could be substantive learning purely from experience, without any feedback. Therefore, in further analysis of the treatment effect on ability I control for experience. Furthermore, it is puzzling that some subjects (16%) performed worse in Part 3. This decrease in performance might be due to subjects following heuristics or making mistakes, leading to unstable performance.

Result 1.2. *In the control group, the experience of playing games in Part 1 (before the irrelevant tutorial) has an insignificant effect on reasoning ability in Part 3 (after the irrelevant tutorial) on average. At the individual level, the effect is mixed and can be substantially large.*

1.5.3 Learning from the Relevant Tutorial

Now I turn to the effect of the relevant tutorial on reasoning ability. As shown in the right panel of Table 1.3, about 56% of subjects in the treatment group have ability

Table 1.3: Subjects' types before and after the tutorials (Counts).

		Control: Irrelevant Tutorial Part 3: After Tutorial						Treatment: Relevant Tutorial Part 3: After Tutorial					
		A0	A1	A2	A3	A4	Total	A0	A1	A2	A3	A4	Total
Part 1: Before Tutorial	A0	16	11	4	1	2	34	13	2	1	1	18	35
	A1	5	16	3	1	0	25	8	10	1	3	6	28
	A2	1	3	3	0	0	7	1	0	0	0	6	7
	A3	0	0	1	1	0	2	0	0	0	0	3	3
	A4	0	0	0	2	7	9	0	0	0	0	7	7
	Total	22	30	11	5	**9**	77	22	12	2	4	**40**	80

Notes: The shades of the cells represents how ability changes from Part 1 to Part 3: light - lower ability; medium - higher ability; dark - same ability.

improvement ($Ak_3 > Ak_1$, $A4_1$ excluded), which is significantly larger than that (32%) in the control group (Fisher's exact test $p = 0.004$), suggesting a positive effect of the relevant tutorial on reasoning ability in addition to experience.

More importantly, a large proportion of subjects (45%) in the treatment group *learn* to play equilibrium and achieve $A4$ after the relevant tutorial ('*learners*' hereafter).[9] The proportion of learners in the control group is just 3%. Again, Fisher's exact test shows a significant difference in the proportions of learners between the two groups ($p = 0.000$).

Moreover, there exist some subjects who do not learn from the relevant tutorial and remain low ability in Part 3. Nevertheless, it is more likely to happen to low-ability (before the tutorial) subjects and perhaps is related to their cognitive ability to learn, which will be discussed in the next section.

[9]Subjects who are classified as $A4_1$ in Part 1 (before the tutorials) are not labeled as learners or non-learners because there is no scope for further improvement in terms of reasoning depths for those subjects. That is, $A4_1$ subjects are labeled as missing value in a dummy variable 'learner' taking the value of 1 if a subject is $A4$ in Part 3, and 0 otherwise. Note that in the control group, two subjects classified as $A4_1$ in Part 1 became $A3_3$ in Part 3. These two subjects are treated as missing value based on the definition of 'learner'. Because the two subjects' ability changed, it might be reasonable to treat them as non-learners, and to label them as 0 rather than missing value. The results are consistent when the two subjects are labelled as either missing value or 0.

Result 1.3. *In the treatment group, 56% of subjects have ability improvement after the relevant tutorial and 45% of them learn to play equilibrium.*

1.5.4 Ability Bound and Cognitive Ability

Now I examine the relationship between ability bounds (before and after the tutorials) and cognitive ability measured by four Cognitive Reflection Test (CRT) questions (Toplak et al., 2011) controlling for demographics including gender, major in economics, age and study level (see Table A.3 for summary statistics of these variables).

In Table 1.4, columns (1)-(3) show that there is no significant group difference in pre-tutorial ability (Ak_1) between the control and treatment group, consistent with the results shown in section 1.5.1. Further, the treatment significantly increases the probability of being a higher type after the tutorials (columns (4)-(5)), as well as learning to play equilibrium (columns (7)-(8)), after controlling for pre-tutorial ability, cognitive ability and demographics.

Turning now to the effect of cognitive ability (CRT) on iterative reasoning ability, CRT significantly increases the probability of being a higher type both before and after the tutorials on average, as well as the probability of learning to play equilibrium. Interestingly, the interaction of treatment and CRT in column (6) has a significantly positive effect on post-tutorial ability, but the treatment and CRT on their own are insignificant in that specification (at 5% level). It suggests that the treatment effect is dependent on CRT. Specifically, the treatment significantly increases reasoning ability for high-cognition subjects, but not low-cognition subjects. Further, the treatment effect increases with CRT, indicating that subjects with higher cognition can learn better and have larger ability improvement after they were taught iterative reasoning. Such heterogeneous treatment effect depending on cognitive ability (column (6)), together with the result that subjects with higher cognition have higher reasoning ability before

Table 1.4: Treatment effect and cognitive ability.

	Ak_1			Ak_3			$A4_3$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	-0.128	-0.171	-0.009	1.053***	0.961**	-1.274*	3.282***	3.295***	1.504
	(0.326)	(0.328)	(0.484)	(0.406)	(0.457)	(0.678)	(0.754)	(0.745)	(2.350)
CRT	0.288***	0.252**	0.287	0.496***	0.464***	0.063	0.712***	0.693***	0.130
	(0.112)	(0.108)	(0.217)	(0.108)	(0.113)	(0.121)	(0.175)	(0.190)	(0.838)
Treatment×CRT			-0.063			0.891***			0.637
			(0.249)			(0.190)			(0.857)
Female		-0.497**	-0.496**		-0.161	-0.185		-0.027	-0.031
		(0.217)	(0.216)		(0.401)	(0.417)		(0.665)	(0.675)
Economics		0.507	0.508		0.709*	0.682*		1.405**	1.416*
		(0.326)	(0.323)		(0.381)	(0.358)		(0.709)	(0.731)
Age		-0.014	-0.014		-0.001	-0.005		-0.007	-0.008
		(0.019)	(0.019)		(0.023)	(0.023)		(0.053)	(0.054)
Study Level		0.199	0.201		0.042	0.038		-0.058	-0.049
		(0.213)	(0.215)		(0.232)	(0.231)		(0.385)	(0.402)
Ak_1				0.701***	0.661***	0.718***	0.021	-0.085	-0.107
				(0.127)	(0.140)	(0.153)	(0.265)	(0.336)	(0.359)
N	157	155	155	157	155	155	141	139	139

Notes: Ak_1 is the pre-tutorial reasoning ability; Ak_3 is the post-tutorial ability; $A4_3$ is an indicator of subjects learning to play equilibria via the tutorials. Session clustered standard error in parenthesis. Two subjects chose to not state their gender. Columns (1)-(6): ordered logit. Columns (7)-(9): logit, $A4_1$ excluded; constant omitted. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the tutorials (columns (1)-(3)), suggest that cognitive ability plays an important role not only in performance but also in the ability to improve performance via training or teaching.[10]

Result 1.4. *More cognitively able subjects are significantly more likely to have higher reasoning ability before the tutorials, as well as more likely to learn better from the relevant tutorial.*

[10] As shown in Table 1.4, there is a significant gender difference in the pre-tutorial ability (columns (2)-(3)), and studying economics has a significantly positive correlation with the post-tutorial ability at 10% (columns (5)-(6)) and with the probability of learning to play equilibrium at 5% (columns (8)-(9)). However, these results are not consistent when using other classification rules (see footnote 7 and 8), thus these results are not discussed in detail here.

1.6 Experimental Results: Self-Assessment

Recall that subjects were incentivised to report their subjective beliefs about their performance in Part 1 (before the tutorials) and in Part 3 (after the tutorials) after the games (see design in section 1.3). These self-assessed performance are designed to study how subjects perceive their performance and improvement in performance after the tutorials (see hypotheses in section 1.4).

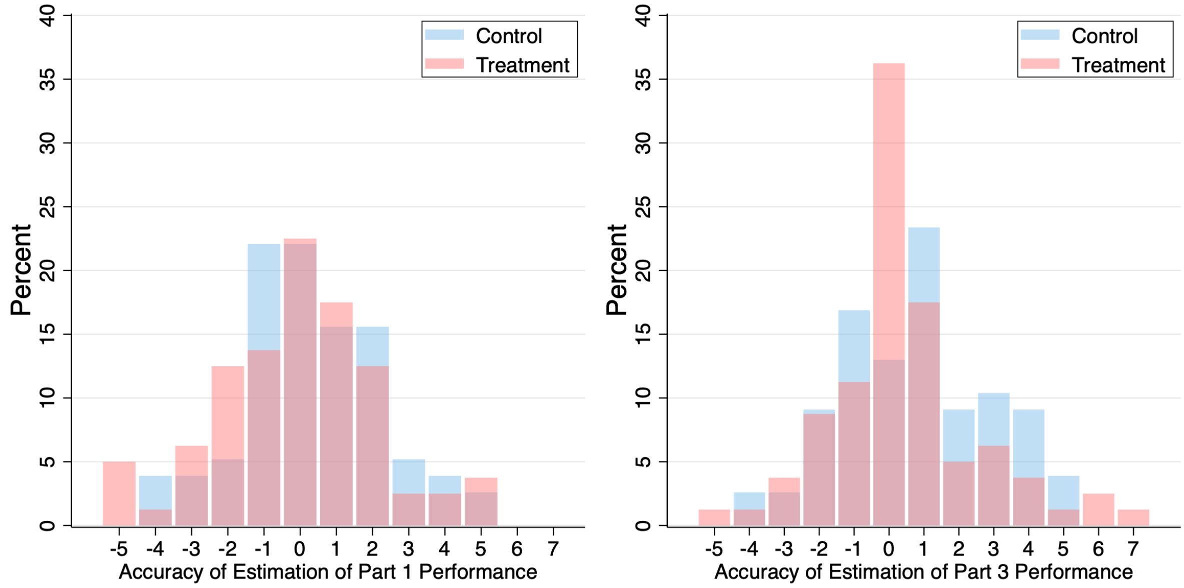
1.6.1 Perceived Performance

Here I first turn to subjects' perception of their own reasoning ability by comparing their estimated performance with actual performance. Figure 1.3 plots histograms of the estimation accuracy, defined by the difference between the estimated performance ($\hat{x}$) and the actual performance (x), across the control and treatment groups.[11] As shown in the left panel regarding estimated pre-tutorial performance, in both groups, only about 22% subjects made the precise estimation, i.e., $\hat{x} - x = 0$, while both underestimation ($\hat{x} - x < 0$) and overestimation ($\hat{x} - x > 0$) of the performance are prevalent. Nevertheless, for the pre-tutorial performance estimations, the treatment group does not behave significantly differently from the control group (Fisher's exact test $p = 0.859$).

In contrast with the pre-tutorial performance estimations, there is a significant group difference in post-tutorial performance estimations (Fisher's exact test $p = 0.003$). In particular, the treatment has a larger proportion of subjects who made precise estimations of their post-tutorial performance (36%), almost triple the proportion of such subjects in

[11] One can define estimation accuracy as the absolute distance between the estimated and actual performance $|\hat{x} - x|$. Table A.8 provides qualitatively the same results in terms of the treatment effect.

(a) $\hat{x}_1 - x_1$. (b) $\hat{x}_3 - x_3$.



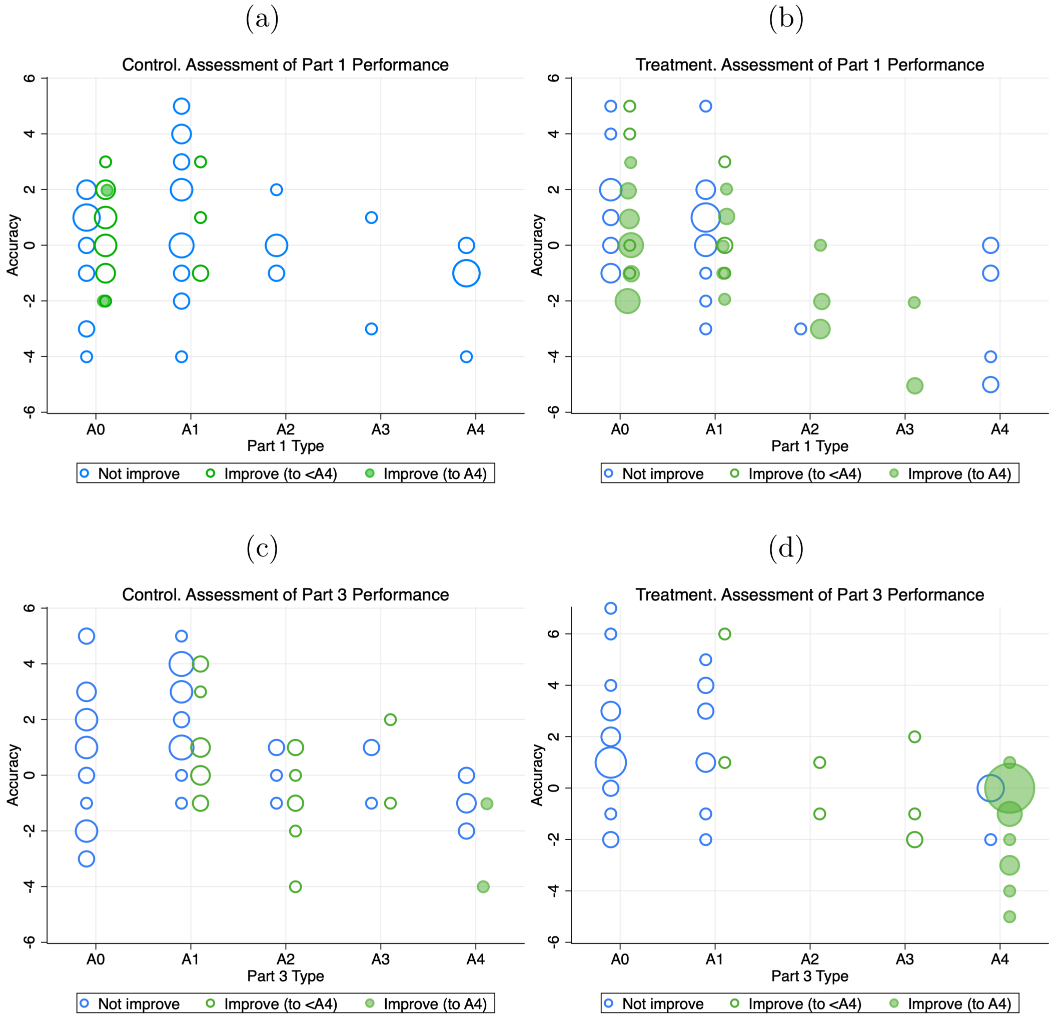
Notes: The estimation accuracy defined by estimated performance − actual performance ($\hat{x} - x$).

Figure. 1.3: Histograms of the estimation accuracy.

the control group (13%). However, underestimation and overestimation of the performance still seem to be common in both groups.[12]

Figure 1.4 shows the estimation accuracy against reasoning ability (Ak). The top panel displays a negative relationship between the pre-tutorial performance estimation and pre-tutorial ability Ak_1. It shows that a lower ability type tends to overestimate while a higher type tends to underestimate, but the higher ability types do not seem to make precise estimations more often. Such tendency does not seem to be corrected by the treatment, and is persistent even for subjects who have ability improvement. A similar relationship can also be seen for the estimations of post-tutorial performance and post-tutorial ability Ak_3 (the bottom panel).

[12]Notice that in the control group, more subjects overestimate and fewer subjects estimate precisely post-tutorial performance compared to pre-tutorial performance. In contrast, in the treatment group, more subjects estimate precisely and fewer subjects under estimate post-tutorial performance compared to pre-tutorial performance. However, marginal homogeneity test suggests no significant difference for both groups ($p > 0.1$).



Notes: The size of the bubbles represents the share of subjects in a category. -0.1 (0.1) is added to each value of Ak to the category 'Not improve' (Improve) to minimize bubbles overlay. Note that subjects have ability $A4$ in Part 1 before the tutorials are placed in the category 'Not improve'.

Figure. 1.4: Assessment accuracy $\hat{x} - x$ against ability type Ak across parts and groups.

Table 1.5: Relationship between estimated pre-tutorial performance, treatment and ability bounds.

	Estimations of pre-tutorial performance					
	(1)			(2)		
	under	precise	over	under	precise	over
Treatment	-0.013	0.004	0.010	-0.005	0.006	-0.001
	(0.056)	(0.081)	(0.078)	(0.059)	(0.084)	(0.081)
Ak_1	0.083**	0.026	-0.115***	0.079**	0.023	-0.108**
	(0.035)	(0.024)	(0.044)	(0.037)	(0.024)	(0.046)
Ak_3	0.051***	0.002	-0.056**	0.039*	0.009	-0.051**
	(0.018)	(0.017)	(0.024)	(0.021)	(0.017)	(0.025)
CRT				0.046	-0.023	-0.024
				(0.032)	(0.022)	(0.034)
Female				0.013	-0.009	-0.004
				(0.070)	(0.051)	(0.063)
Economics				-0.041	-0.029	0.074
				(0.068)	(0.066)	(0.080)
Age				-0.004	-0.004	0.009**
				(0.006)	(0.007)	(0.004)
Study Level				0.070*	0.002	-0.076
				(0.036)	(0.047)	(0.068)
N	459	459	459	454	454	454

Notes: Average marginal effects from mixed logit regressions (session clustered standard error used) of estimation of pre-tutorial performance. Standard errors of marginal effects in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Both the group difference and the correlation between estimated performance and Ak are confirmed by the regression results shown in specifications (1) and (2) of Tables 1.5 and 1.6, which report the average marginal effect from mixed logit regressions of the three types of estimated performance.[13] I find that the treatment effect on pre- and post-tutorial performance estimates differs. The treatment does not significantly affect the pre-tutorial performance estimates. One possible explanation of this result could be that the subjects in the treatment group erroneously believe that their methods used

[13]By using mixed logit, I account for the situations that subjects at two endpoints of the scale (i.e., the actual performance is 0 or 8) cannot underestimate or overestimate.

to derive payoff-maximizing decisions were equivalent to iterative dominance reasoning. However, they were unaware that their methods would turn out to be non-optimal.

In contrast, the treatment decreases the probability of underestimation of the post-tutorial performance while increases the probability of precise estimations. Note that the vast majority of the precise estimations of the post-tutorial performance in the treatment group are made by the subjects who learned to play equilibrium (i.e., learners) (see Figure 1.4d). Thus, it is possible that the treatment effect on the estimations of the post-tutorial performance is driven by the learners. Specification (3) of Table 1.6 shows that indeed the treatment does not significantly affect estimations among non-learners.[14]

Overall, the subjects who have low ability before the tutorials tend to overestimate their pre-tutorial performance, but would make more accurate post-tutorial performance estimations if they learn to play equilibrium via the relevant tutorial. That is, the estimation accuracy enhancement is accompanied by the learning outcomes: subjects who learn better understand themselves better. However, if they do not learn, they tend to remain low ability but falsely believe they have always made the right decisions.

Result 1.5. *A lower type tends to overestimate while a higher type tends to underestimate. The treatment does not correct such tendencies for subjects who do not learn to play equilibria.*

1.6.2 Perceived Performance Improvement

In this subsection, I examine the perceived performance improvement using the difference between estimates of pre- and post-tutorial performance. As subjects were asked to estimate performance in the two parts on the same page, it was a direct question to them

[14] Due to data spareness (i.e., all of the learners in the control group underestimate), I am unable to use the interaction term $A4_3\times$ Treatment to test whether the treatment effect on estimation accuracy is channeled through learning outcome. Thus I omit learners, and re-run the regressions.

Table 1.6: Relationship between estimated post-tutorial performance, treatment and ability bounds.

	Estimations of post-tutorial performance								
	(1)			(2)			(3)		
	under	precise	over	under	precise	over	under	precise	over
Treatment	-0.180***	0.133***	0.066	-0.172***	0.130***	0.058	-0.003	-0.064	0.067
	(0.054)	(0.049)	(0.062)	(0.055)	(0.046)	(0.052)	(0.064)	(0.043)	(0.063)
Ak_1	-0.082**	0.022	0.085	-0.066*	0.004	0.087	-0.085	0.000	0.085
	(0.037)	(0.020)	(0.052)	(0.039)	(0.023)	(0.056)	(0.056)	(0.043)	(0.071)
Ak_3	0.051**	0.035*	-0.120***	0.058*	0.016	-0.103**	0.114***	-0.013	-0.101*
	(0.025)	(0.020)	(0.046)	(0.033)	(0.021)	(0.048)	(0.042)	(0.035)	(0.055)
CRT				0.005	0.044	-0.069**	0.048*	0.020	-0.068**
				(0.026)	(0.039)	(0.033)	(0.027)	(0.023)	(0.029)
Female				-0.029	-0.011	0.057	0.020	-0.062	0.042
				(0.080)	(0.061)	(0.076)	(0.108)	(0.075)	(0.091)
Economics				-0.234**	0.063	0.239**	-0.254*	0.017	0.238*
				(0.093)	(0.058)	(0.111)	(0.131)	(0.057)	(0.124)
Age				0.001	-0.012	0.015*	0.006	-0.044**	0.038***
				(0.007)	(0.009)	(0.009)	(0.008)	(0.019)	(0.013)
Study Level				0.010	0.023	-0.045	-0.017	0.132***	-0.115**
				(0.048)	(0.056)	(0.054)	(0.043)	(0.040)	(0.058)
N	426	426	426	421	421	421	315	315	315

Notes: Average marginal effect from mixed logit regressions (session clustered standard error used) of estimations of post-tutorial performance. Standard errors of marginal effects in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Learners are excluded in specification (3).

if they believe they have improved or not. Thus, if subjects believe they did improve, they would give a larger estimate of post-tutorial performance than pre-tutorial performance.

As reported in the bottom row of Table 1.7, subjects in both control and treatment groups give a larger number for Part 3 than Part 1: the mean (standard deviation) of the difference between two estimations ($\hat{x}_3 - \hat{x}_1$) is 0.636 (1.708) for the control group and 2.463 (2.261) for the treatment group. Both means are significantly larger than zero (one-sided t-test $p < 0.01$), suggesting that subjects in both groups believed they performed better in Part 3 than Part 1 on average.

Note that this could result from experience and/or the relevant tutorial for the treatment group. To see if the relevant tutorial also plays a role in the perceived improvement, one needs to separate the two components, and further look at if the treatment group believes they have greater improvement than the control group. Indeed the treatment group is found to have a significantly larger mean than the control group (one-sided t-test $p < 0.01$), thus suggesting that the perceived improvement in the treatment group could be attributed to the relevant tutorial in addition to experience. Finally, the regression results reported in columns (1)-(2) of Table 1.8 also show a significant group difference in the perceived performance improvement (a dummy for $\hat{x}_3 > \hat{x}_1$), after the ability improvement (a dummy for $Ak_3 > Ak_1$), CRT and demographics are additionally controlled for.

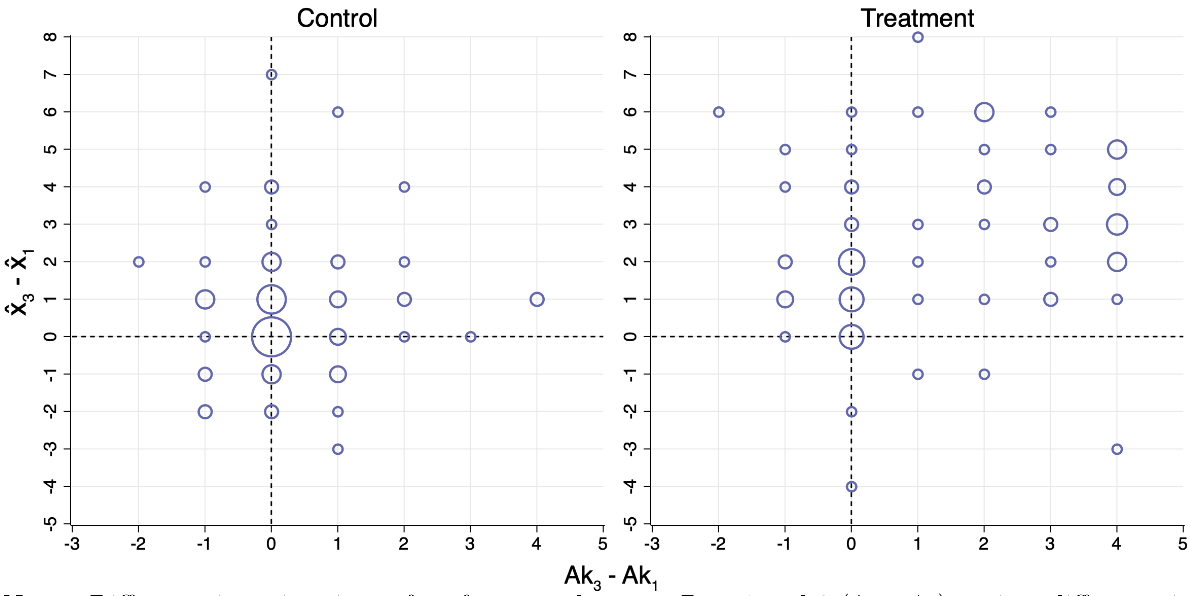
Table 1.7: Perceived improvement vs. actual improvement.

$\hat{x}_3 - \hat{x}_1$	Control	Treatment	Difference
$Ak_3 \leq Ak_1$	0.630***	1.469***	0.838**
	(1.768)	(1.951)	
$Ak_3 > Ak_1$	0.682*	3.244***	2.562***
	(1.912)	(2.234)	
Total	0.636***	2.463***	1.826***
	(1.708)	(2.261)	

Notes: Mean (SD) of the difference in estimations of performance between Part 1 and Part 3 ($\hat{x}_3 - \hat{x}_1$) of each category. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ if a value is significantly different from zero according to one-sided t test.

1.6.2.1 Actual Improvement vs. Perceived Improvement

Here I show that both subjects who exhibit the ability improvement and those who do not believe that they have ability improvement. Figure 1.5 is a bubble plot of the estimation difference $\hat{x}_3 - \hat{x}_1$ against the reasoning ability difference $Ak_3 - Ak_1$ between the two parts. It is clear that the treatment group tends to believe they performed better in Part 3 than in Part 1, i.e., $\hat{x}_{3T} > \hat{x}_{1T}$ regardless of their ability improvement. Such tendency is shown significant at 1% level for subjects in the treatment group who have ability



Notes: Difference in estimations of performance between Part 1 and 3 ($\hat{x}_3 - \hat{x}_1$) against difference in reasoning ability between two parts ($Ak_3 - Ak_1$). Bubble size is proportional to the share of subjects. The horizontal dashed reference line: $\hat{x}_3 = \hat{x}_1$; the vertical dash reference line: $Ak_3 = Ak_1$.

Figure. 1.5: Perceived improvement vs. actual improvement.

improvement and for those who do not (see Table 1.7). Further, for both subgroups, treatment group obtains a significantly larger mean of $\hat{x}_3 - \hat{x}_1$ than the control group (one-sided t-test $p < 0.05$).

Columns (3)-(4) of Table 1.8 show that the treatment effect remains significantly positive after adding interaction term treatment$\times(Ak_3 > Ak_1)$. It confirms that subjects who have no improvement also perceived the relevant tutorial to be useful in addition to experience.

Result 1.6. *Both those subjects who have ability improvement and those who do not believe that they have ability improvement from the relevant tutorial.*

As discussed in the introduction, the perceived improvement could be interpreted as the perceived effectiveness of experience and/or the relevant tutorial in improving performance: subjects in the control group would believe they perform better only when they deem that the introspection from experience made them understand the

Table 1.8: Test of the perceived performance improvement.

$\hat{x}_3 > \hat{x}_1$	(1)	(2)	(3)	(4)
Treatment	1.543***	1.583***	1.113***	1.163***
	(0.419)	(0.400)	(0.416)	(0.390)
$Ak_3 > Ak_1$	0.826**	0.793*	0.357	0.341
	(0.386)	(0.444)	(0.432)	(0.473)
Treatment×($Ak_3 > Ak_1$)			1.244	1.244
			(0.760)	(0.852)
CRT		0.088		0.038
		(0.169)		(0.176)
Female		-0.165		-0.182
		(0.358)		(0.375)
Economics		-0.134		-0.219
		(0.309)		(0.336)
Study Level		0.139		0.147
		(0.264)		(0.261)
Age		-0.012		-0.017
		(0.016)		(0.016)
Constant	-0.325	-0.456	-0.174	-0.063
	(0.248)	(0.799)	(0.238)	(0.930)
N	141	139	141	139

Notes: Logit. $\hat{x}_3 > \hat{x}_1$ is an indicator for subjects estimating a larger number of rounds that they made optimal choices in Part 3 than Part 1. $Ak_3 > Ak_1$ is an indicator for subjects with ability improvement. Session clustered standard error in the parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $A4_1$ excluded.

games better. Meanwhile, subjects in the treatment group would have similar behavior when experience *and/or* the relevant tutorial are perceived as effective in performance improvement. By showing that the treatment group believes they have improved more than the control group, one could interpret that the teaching is perceived to be effective by subjects in improving performance.

Note that it is possible that the relevant tutorial could make subjects aware of mistakes made in Part 1, leading to $\hat{x}_{1T} < \hat{x}_{1C}$. Hence, the positive treatment effect on $\hat{x}_3 > \hat{x}_1$ could be driven by the awareness of mistakes rather than by the perceived ability improvement. Using the same specifications in Table 1.8 but replacing the dependent variable with $\hat{x}_1$ or $\hat{x}_3$, I find that the treatment indeed decreases $\hat{x}_1$, however insignificantly (columns (1)-(3) of Table A.7). In contrast, the treatment increases $\hat{x}_3$ significantly at 1% level (columns (4)-(6) of Table A.7). Thus, the error correction mechanism of the treatment *alone* cannot explain the treatment effect on $\hat{x}_3 > \hat{x}_1$, allowing one to interpret $\hat{x}_3 > \hat{x}_1$ as the perceived ability improvement or perceived teaching effectiveness.

Finally, recall that Result 1.5 suggests that subjects who do not learn to play equilibrium do not enhance accuracy of estimations, in particular those subjects who do not improve and persistently have low ability (both before and after the tutorials) keep overestimating their performance, i.e., *overconfidence in the absolute level of performance* of Part 1 and 3. Meanwhile, Result 1.6 shows that subjects who do not improve believe that they do, i.e., *overconfidence in learning ability*. That is, there is evidence of *multi-dimensional overconfidence* of low-ability subjects (also see Figure A.1). The negative correlation between ability/learning outcome and confidence suggests that overconfidence might be detrimental to learning outcomes. It is worth noting that one should be cautious about making inferences from this correlation about *causal* relationship between overconfidence and learning, because the estimated performance was only elicited once in the end of the experiment, namely there is no exogenous variation in the level of overconfidence. Nevertheless, to foster learning for the low-ability subjects, it is possible that a more intensive treatment on strategic thinking or an intervention about overconfidence would be beneficial.

1.7 Conclusion

In this paper, I investigate to what extend subjects' ability to do iterative reasoning is bounded. I also explore whether subjects can be taught to do iterative reasoning, and how subjects perceived their performance and improvement in performance. In the experiment, subjects played the 4-player ring games *against known rational computer players*, before and after either the *relevant* (treatment) or *irrelevant* (control) tutorial, followed by their estimations of the performance in the games.

I find that before the tutorials a majority of subjects (90%) cannot play equilibrium against rational computer players, and thus are ability bounded. Further, more than half of the subjects performed less than 2 steps of iterative reasoning. These low-ability subjects are likely to have difficulty in forming beliefs about opponents' choices, as pointed out by Bosch-Rosa and Meissner (2020). Therefore, the bounded ability is an important aspect of non-equilibrium play.

Second, the relevant tutorial is found to have a strong positive effect on the ability bounds. In particular, 45% of subjects in the treatment group learn to play the equilibrium. Thus, teaching strategic thinking can be highly effective in nudging people to make better strategic decisions even in a short-duration and low-cost lab experiment. It further suggests that teaching economics in a general educational context may have encouraging outcomes, and providing economics courses to non-economics students or even the general population could achieve better decision-making. However, to have any effective teaching or performance improvement, more educational resources might be required for low-cognition students, who are found to have lower reasoning ability before the tutorials and also less likely to learn from the relevant tutorial. Such heterogeneous treatment effect points to the need for the optimization of policy or intervention allocation (Zhou et al., 2022), i.e., identification of group characteristics, and design of treatment intensity specific to each group. Meanwhile, I am cautious about whether this one-off treatment have long-lasting effect on subjects' reasoning ability, because effects may fade away (e.g., Fan, 2000; Alan and Ertac, 2018; Alan et al., 2019; Cappelen et al., 2020). Nevertheless, the education of strategic thinking might have prolonged effect on decision making if it can be integrated since schools.

Finally, I find that low-ability subjects persistently overestimate their performance if subjects did not learn to play equilibrium via the relevant tutorial. I further find that those subjects who do not improve reasoning ability believe they do. In other words, I find that low-ability subjects are overconfident in two dimensions - the absolute level of performance and the ability to improve performance. Fischer and Sliwka (2018) show that overconfidence in performance could be detrimental to learning outcomes while overconfidence in learning ability might not. However, the current paper cannot answer a question whether overconfidence is the *cause* of these subjects not learning from the relevant tutorial, because subjects only estimated performance after the games. Thus, it is of interest for future research to study the causal relationship between the multi-dimensional (over)confidence and learning outcomes. Answers to this question are also of fundamental importance to policy-making. For example, to foster learning, will intervention in overconfidence be effective? Which dimension of overconfidence should the policy be focused on?

To conclude, this paper finds that most of the subjects in general do not think strategically but they can be taught to do so through a simple intervention, the relevant tutorial. Therefore, implementation of a new policy and institutional design ought to be accompanied by some training on the underlying mechanism. On the other hand, low-ability subjects do not learn. To improve their decisions, one might need to implement either a more intensive treatment on strategic thinking or an intervention on overconfidence, which could be a direction of future research.

Chapter 2

Is Level-k Behavior Driven by Ability or Belief? An Experimental Study

2.1 Introduction

Non-equilibrium play or bounded rationality is a well-documented behavior in various strategic settings. Level-k model (Nagel, 1995; Stahl and Wilson, 1994, 1995; Ho et al., 1998) is a prominent model, proposed to explain such behavior by using non-equilibrium beliefs. The model assumes that $L0$ players are irrational and play randomly or according to salient rules. $L1$ players believe all the opponents are $L0$ and best respond to the strategies of $L0$ players. Generally, for $k \geq 1$, a Lk player best responds to $Lk-1$ players' strategy since they believe all the other players are $Lk - 1$.[1]

Most subjects in the level-k literature are found to be no higher than $L4$ (see Crawford et al., 2013, for a review). Yet, it is still unclear whether the observed low levels are driven by low-order beliefs about opponents' rationalities as the theory assumes, or whether they are driven by limited intrinsic reasoning ability. In order to identify the driving factors behind the level-k behavior, one has to disentangle ability and belief from subjects' decisions.

[1]Cognitive hierarchy model (Camerer et al., 2004) is another pioneer work to explain the bounded rationality. It assumes that Lk players best respond to their beliefs of a distribution of all the lower level players, instead of only players of one level below, i.e., $Lk - 1$.

This study involves an experimental design with two novel features, which help to examine the extent to which the rationality levels are affected by ability vs. belief at the individual level. First, subjects played 4-player dominance-solvable ring games (introduced by Kneeland (2015)) *against themselves*. As the result, strategic uncertainty is entirely removed, allowing to measure ability bound (Ak) as playing such games only requires the ability to perform k steps of iterative reasoning. At the same time, subjects played the *same* games on the *same* interface against other human subjects. This allows to measure rationality bound (Lk), as strategic uncertainty is present, and subjects need to form beliefs about opponents' moves. The employed within-subject design is different from the typical within-subject design, as here subjects did not perform the two tasks sequentially but simultaneously, so the order effect is diminished by design. Instead, in effect subjects faced a question of whether to change their action choices or not due to different opponents. This novel design allows one to identify (1) the entire empirical distribution of Ak; (2) the driving factor of the observed Lk behavior: if a subject's level is bounded by her ability, she would display the same reasoning depth facing either of the two types of opponents ($Lk = Ak$); otherwise, the subject is expected to display higher reasoning depth when playing against herself than against other subjects ($Lk < Ak$). The second novel design is that the payoff tables used in this experiment are carefully chosen so that the expected payoff of performing one more step is always higher than stopping at the current level. Further, no action combination beats Nash equilibria. Thus, subjects cannot 'collude' with themselves, and cannot get a higher payoff than playing Nash equilibria when playing against themselves - the games are *collusion-proof*, and subjects were provided incentives to reason until reaching their capabilities.

The experiment was conducted on Amazon Mechanical Turk (Mturk) online. I find that only about 2% of the subjects played equilibria, so a vast majority of the subjects have limited reasoning ability. Additionally, I find that 76% of them performed the same reasoning depth in the two tasks and thus their levels are bounded by limited ability (or Lka). 10% have ability bounds larger than rationality bounds and thus their levels are driven by beliefs (or Lkb). These results suggest that limited ability plays a major role in bounded rationality. Moreover, I find that cognition measured by scores obtained from Cognitive Reflection Test (CRT, Frederick, 2005; Toplak et al., 2011) predicts ability bounds but not rationality bounds, and subjects with higher CRT scores are more likely to be Lkb players.

While the existing literature of bounded rationality focuses on testing whether subjects hold non-equilibrium beliefs (e.g., Slonim, 2005; Costa-Gomes and Crawford, 2006; Palacios-Huerta and Volij, 2009; Agranov et al., 2012; Georganas et al., 2015; Gill and Prowse, 2016), several recent works have explored limited reasoning ability as a determinant of level-k behavior (Alaoui and Penta, 2016; Alaoui et al., 2020; Jin, 2021; Friedenberg et al., 2021; Halevy et al., 2021). The most relevant work is Puente (2020) where subjects played sequential ring games against rational computer players and other human subjects in the standard within-subject setting. Nevertheless, he finds a significant order effect which hinders clear identification of the individual-level driving factor behind bounded rationality. Another closely related paper is Bosch-Rosa and Meissner (2020) where subjects played two-player guessing game against themselves (or one-player game in effect). They find that about 70% of their subjects cannot play equilibrium and have bounded reasoning ability. It is important to note that while one-player guessing game allows for a diagnostic test of whether subjects' reasoning ability is binding, it does not allow for the measurement of exact reasoning steps, in contrast to the ring games used in this paper, which allows for measurement up to 4 steps.

This paper also relates to the literature on the relationship between cognition and strategic sophistication. Gill and Prowse (2016); Burnham et al. (2009); Schnusenberg and Gallo (2011); Capra (2019); Jin (2021) show that subjects with higher cognition tend to have higher rationality bounds. Brañas-Garza et al. (2012) finds that a positive association between rationality bound with CRT score but not with Raven's test score. However, Georganas et al. (2015); Sprenger and Zhao (2021) do not find evidence of such relationship between levels and CRT. This paper contributes to the literature by finding that CRT scores predict ability bounds but not rationality bounds. Further I find that *Lkb* subjects have significantly higher CRT scores than *Lka* subjects.

The remainder of the paper is structured as follows: Section 2.2 describes the ring games and empirical identification strategy of ability and rationality bounds. Section 2.3 and 2.4 provide experimental design and procedure. Section 2.5 reports results. Finally, section 2.6 concludes.

2.2 Ring Game and Identification Strategy

The top panel of Figure 2.1 displays a 4-player simultaneous ring game with three actions (Kneeland, 2015). The game has a specific structure where a player's payoff directly depends only on the action of herself and one adjacent player in the ring: Player 1's payoff depends on Player 1 and Player 2's decisions; Player 2's payoff depends on Player 2 and Player 3's decisions; Player 3's payoff depends on Player 3 and Player 4's decisions; Player 4's payoff depends on Player 4 and Player 1's decisions. Importantly, there is a key player, i.e., Player 1 who has a dominant strategy. Thus, every other player's best response can be computed iteratively.[2]

Therefore, the game is dominance solvable, and the type of a subject (denoted by $Tk \in \{Lk, Ak\}$) is determined by the iterative reasoning steps that the subject displays with and without strategic uncertainty respectively, i.e., playing games against other

[2] In contrast to Kneeland (2015) (and Chapter 1) where the key player is Player 4, this experiment has Player 1 as the key play. In the pilot session where the key player is randomly assigned, subjects facing Player 1 as the key player tend to perform better. Thus, this design provides an upper bound of reasoning ability.

G1

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f	Sum
a	14	6	16	(36)
b	26	22	18	(66)
c	18	4	14	(36)

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i	Sum
d	12	24	16	(52)
e	18	14	28	(60)
f	26	8	18	(52)

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l	Sum
g	26	14	6	(46)
h	8	10	28	(46)
i	12	30	14	(56)

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c	Sum
j	32	4	6	(42)
k	6	2	36	(44)
l	4	26	12	(42)

G2

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f	Sum
a	26	22	18	(66)
b	18	4	14	(36)
c	14	6	16	(36)

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i	Sum
d	12	24	16	(52)
e	18	14	28	(60)
f	26	8	18	(52)

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l	Sum
g	26	14	6	(46)
h	8	10	28	(46)
i	12	30	14	(56)

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c	Sum
j	32	4	6	(42)
k	6	2	36	(44)
l	4	26	12	(42)

Notes: The best responses are in red and bold. Payoff sum of each action is in the parenthesis, with the highest sum (i.e., risk dominance action) underlined.

Figure. 2.1: The Ring Games.

human subjects and themselves respectively. For example, a $T4$ subject would play the (iterated) best responses in every player position, while a $T3$ subject would do so when she is at the position of Player 1, 3, and 4, but not Player 2 where getting the iterated best response would require 4 reasoning steps.

Importantly, a $T3$ subject might choose the iterated best response by chance at Player 2's position, leading this $T3$ subject to be misidentified as a $T4$ subject. Following Kneeland (2015), this issue is resolved by the introduction of the second ring game (bottom panel of Figure 2.1). The two games are identical, except that the dominant strategy of Player 1 in the second game has a different label, resulting in a different iterated best response of every other player. Additionally with Kneeland (2015)'s Natural Exclusion Restriction, now it is possible to separate high type subjects from low type subjects ($k \geq 1$): due to the same payoff tables, a subject is assumed to play the same

Table 2.1: Predicted action profiles for individuals with different levels and abilities.

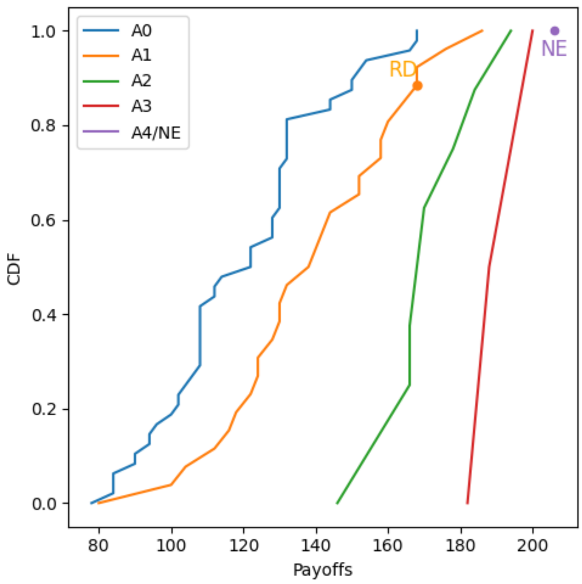
Lk/Ak	Player 1	Player 2	Player 3	Player 4
$L4/A4$	(b,a)	(d,f)	(h,g)	(l,j)
$L3/A3$	(b,a)	$(d,d)(e,e)(f,f)$	(h,g)	(l,j)
$L2/A2$	(b,a)	$(d,d)(e,e)(f,f)$	$(h,h)(i,i)(g,g)$	(l,j)
$L1/A1$	(b,a)	$(d,d)(e,e)(f,f)$	$(g,g)(h,h)(i,i)$	$(j,j)(k,k)(l,l)$

Notes: The first entry is the action in G1 while the second entry is the action in G2.

action in the paired games at a player position if the required reasoning depth is beyond her ability or rationality bound. For example, at Player 2's position, the $T3$ players will choose (d,d), (e,e), or (f,f) in G1 and G2 respectively, which requires four steps of iterative reasoning, above their limits. Similarly, one can get the predicted action sets at each player position of every type of subjects, which are presented in Table 2.1.

Moreover, at each player position except the key player position, the payoff matrices in this experiment are designed to separate actions derived using payoff dominance strategy from those relying on heuristics such as maximax strategy or risk dominance strategy (RD i.e., choosing action gives the highest sum of payoffs across columns, underlined in Figure 2.1). This design addresses the misidentification problem due to action coincidence of different strategies (Jin, 2022; Sprenger and Zhao, 2021).

Furthermore, Ak players' payoffs of all possible action profiles are designed to first order stochastic dominate $Ak-1$ players' payoffs, as illustrated in Figure 2.2. Thus, subjects are incentivized to reason till reaching their ability bounds. More importantly, Nash equilibria are designed to have the highest payoff compared to *any* combination of action choices. It makes the games *collusion proof* - subjects cannot collude with themselves to get a higher payoff by playing some actions than playing Nash equilibrium when they play the games against themselves.



Notes: It is assumed that the payoffs of all the predicted action profiles given Ak have equal probability.

Figure. 2.2: CDF of payoff of each ability bound (G1 & G2).

2.3 Experimental Design

Subjects first completed a survey on demographics and cognitive ability measured by Cognitive Reflection Test (Frederick, 2005; Toplak et al., 2011). Afterwards, they read the instructions, and had to pass the comprehension quiz before proceeding with the games.[3]

Subjects faced two tasks *simultaneously* on the same page (see interface screenshot in Appendix B.3) for two rounds. Subjects played one of the paired ring games in each round (as shown in Figure 2.1). The only difference between the two tasks was the opponents. Specifically, at each player position, subjects played against 3 other human subjects in Task 1, and played the same game against themselves in Task 2. Thus, subjects faced a direct question of whether to choose different actions or not due to different opponents. Moreover, it mediates the learning effect or order effect compared to

[3] See Appendix B.2 for the instructions.

standard within-subject design. No feedback was provided, and one out of eight decisions from either task was randomly selected for payment to avoid hedging.

With different opponents, Task 1 (against others) measures rationality bounds, while Task 2 (against oneself) measures ability bounds without any strategic uncertainty. A player whose level is bounded by her reasoning ability would perform the same reasoning steps in the two tasks. In contrast, if the player's level is bounded by belief instead of ability, she would increase reasoning steps when the belief component is totally removed in Task 2. A Lk player is called ability-bounded Lk (Lka) if $Ak = Lk$, and belief-bounded Lk (Lkb) if $Ak > Lk$. Note that since in 4-player ring games, the maximum step to solve the games is four, one cannot identify determinants for $L4$ players.

2.4 Experimental Procedures

The experiment was conducted on Amazon Mechanical Turk (Mturk) in 2020 using O-Tree (Chen et al., 2016). The Mturk subjects are well known for attrition, so there might be insuffcient subjects to form 4-player groups. To address the issue, I randomly matched subjects *after* a session finished. When some subjects were left unmatched, I randomly re-selected subjects who have already had a group into the incomplete group. The payoffs of the re-selected subjects were not altered. The matching was anonymous and fixed across the two rounds. Because of ex-post matching, subjects only knew about their payoffs when they received the earnings. Subjects were informed of this procedure.

To participate in our experiment, Mturk subjects have to be residents in the US, and have at least 100 approved assignments and 95% approval rate.[4] 179 subjects completed the whole experiment (and 148 subjects drop out). Subjects actively spent 16.12 minutes

[4]See Hauser et al. (2019) for the discussion on subjects' qualifications required on Mturk.

on the experiment on average.[5] The average earnings were \$4.31 (including \$2.50 reward for accepting the experiment and completing the survey). Subjects who drop out were not paid. See Table B.5 for summary statistics of the sample.

2.5 Results

2.5.1 Ability *vs.* Belief

Following the identification strategy discussed in section 2.1, a subject is assigned a type ($Tk \in \{Ak,\ Lk\}$) if her action profile is matched with the type's predicted action profile allowing for up to one deviation from the predictions. If she can be identified as more than one type with the deviation, she is assigned to the lower type conservatively.[6,7] Finally, if subjects made more than one deviation from the predicted profiles, they are assigned to $T0$.[8]

As shown in Table 2.2, most subjects have limited iterative reasoning ability - only 2% of my subjects played equilibria when playing against themselves and are identified as $A4$. 2% of the subjects are $A3$, 4% are $A2$, 22% are $A1$, 70% are $A0$.

Result 2.1. *When playing against themselves, about 2% of the subjects played Nash equilibrium and are* $A4$*, and the remaining 98% have ability lower than* $A4$*, and thus are limited ability.*

[5]If subjects are inactive such as not moving a mouse or hitting a keyboard for more than 30 seconds, switching the experiment tab away, or minimizing the window, an invisible timer stops. Including inactivities, subjects spent 20.92 minutes on average.

[6]Consider an action profile $(b,\ a)$, $(d,\ \underline{e})$, $(h,\ g)$, $(l,\ j)$ where the mistake is underlined. She might be a $T4$ player if the mistake is a deviation from f to e, while she might also be a $T3$ player if the mistake is a deviation from d to e. Conservatively, she will be classified as $T3$.

[7]Table B.1 shows type classification if subjects are assigned to the higher type when they can be identified as two types with an error. All the results presented below are qualitatively the same if this classification rule is used.

[8]Table B.2 reports subjects' types using a classification rule with two deviations. Allowing for two mistakes does not alter the results reported below.

Table 2.2: Abilities and levels.

		Against Others					Total	
		L0	L1	L2	L3	L4		
Against Yourself	A0	103	18	5	0	0	126	70.39%
	A1	9	30	1	0	0	40	22.35%
	A2	3	2	2	0	0	7	3.91%
	A3	0	1	1	1	0	3	1.68%
	A4	1	0	0	0	2	3	1.68%
	Total	116	51	9	1	2	179	
		64.80%	28.49%	5.03%	0.56%	1.12%		

Notes: Percentages in the rightmost column and the bottom row. Counts in the table cells. The shades of the cells represent the changes in reasoning depth: light - $Ak > Lk$; medium - $Ak < Lk$; dark - $Ak = Lk$.

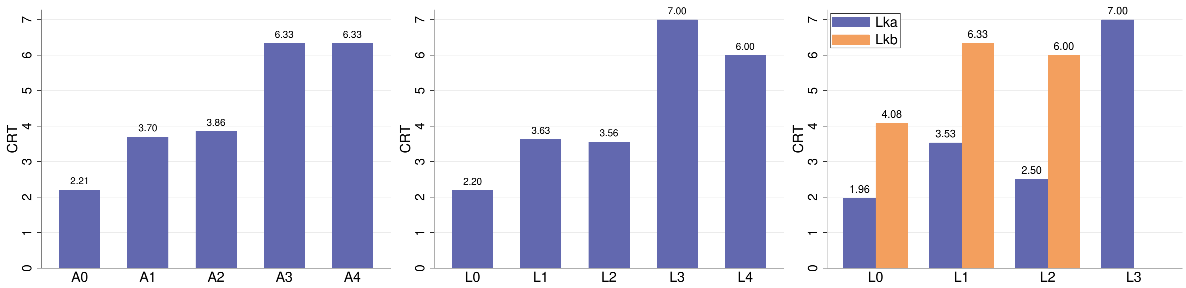
As aforementioned, a belief-bounded Lk player (Lkb) would have $Ak > Lk$ while an ability-bounded Lk player (Lka) would have $Ak = Lk$. The change in reasoning depth in the two tasks is illustrated in Table 2.2. Excluding $L4$, 76% of the subjects show the same reasoning depth when facing different opponents, and thus are Lka. On the other hand, 10% have larger ability bounds than rationality bounds, and thus are Lkb. However, it is puzzling that the remaining 12% have larger rationality bounds than ability bounds.

Result 2.2. *About 76% of the subjects have the same reasoning depth across the two tasks thus are ability-bounded Lk players, while about 10% increase their reasoning depth when playing against themselves compared to others, thus are belief-driven Lk players.*

2.5.2 CRT

This subsection looks at the relationship between subjects' type and cognitive ability measured by the total score of seven questions of the Cognitive Reflection Test (Frederick, 2005; Toplak et al., 2011). The subjects obtained a mean CRT score of 2.74 with a standard deviation of 2.40 and a median of 2.

(a) CRT against Ak (b) CRT against Lk (c) CRT against Lka and Lkb



Notes: In panel (c) $L3b$ is empty due to zero observation; $L4$ is empty due to un-determined driving factor as noted earlier in section 2.3.

Figure. 2.3: Relationship between CRT and type.

Panel (a) of Figure 2.3 shows that CRT is positively associated with ability bounds. The ordered logit regression results in column (1) of Table 2.3 confirm that this positive relationship is significant at 1% level, after controlling for subjects' characteristics i.e., age, gender, study level, major, risk preference, and experience of games measured by log value of the number of assignments subjects have completed on Mturk, using robust standard error. This relationship between CRT and Ak has also been seen in Chapter 1 (Result 1.4).

Interestingly, CRT also significantly increases the probability of being Lkb player when Lk and demographics are controlled (Panel (c) of Figure 2.3 and columns (5)-(7) of Table 2.3). This is consistent with Jin (2021) who finds that $L2b$ subjects have significantly higher CRT scores than $L2a$ subjects. It could be explained by the positive relationship between Ak and CRT: Lkb subjects have higher reasoning ability than Lka subjects, and thus on average are more likely to possess higher CRT scores.[9]

Furthermore, I find that CRT has a significant positive correlation with rationality bound Lk when only controlling for demographics (panel (b) of Figure 2.3 and column (2) of Table 2.3). However, when I further control for Ak, CRT on its own no longer has a significant effect (column (3)). Thus, it is possible that CRT affects Lk through Ak,

[9]When using different classification rules (see footnote 7 and 8), the results regarding CRT are consistent except column (7) of Table 2.3, so this specification is not discussed in details here.

Table 2.3: Relationship between subjects' types and CRT.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ak		*Lk*			*Lkb*	
CRT	0.348***	0.205**	0.037	0.151	0.511***	0.534***	0.472***
	(0.098)	(0.091)	(0.084)	(0.124)	(0.131)	(0.131)	(0.132)
Age	-0.013	-0.001	0.000	-0.000	-0.020	-0.019	-0.025
	(0.019)	(0.015)	(0.014)	(0.015)	(0.028)	(0.027)	(0.028)
Study Level	0.546*	0.352	0.178	0.208	-0.285	-0.206	-0.284
	(0.296)	(0.287)	(0.300)	(0.306)	(0.552)	(0.581)	(0.610)
Female	-0.218	-0.443	-0.304	-0.409	0.436	0.346	0.414
	(0.421)	(0.403)	(0.407)	(0.428)	(0.655)	(0.665)	(0.681)
Log Hits	0.108	0.127	0.115	0.107	-0.027	-0.014	0.012
	(0.099)	(0.081)	(0.082)	(0.091)	(0.152)	(0.156)	(0.158)
Science & Economics	0.243	-0.171	-0.168	-0.270	-0.161	-0.229	-0.225
	(0.389)	(0.351)	(0.367)	(0.396)	(0.638)	(0.619)	(0.627)
Risk	-0.061	-0.108	-0.115	-0.121	0.068	0.068	0.071
	(0.088)	(0.085)	(0.083)	(0.088)	(0.139)	(0.138)	(0.138)
Ak			1.558***	2.520***			
			(0.416)	(0.725)			
Ak×CRT				-0.211			
				(0.161)			
Lk						-0.425	-2.511*
						(0.508)	(1.283)
Lk×CRT							0.356*
							(0.207)
N	143	143	143	143	122	122	122

Notes: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns (1)-(4): ordered logit; Columns (5)-(7): logit; undetermined subjects (Lk >Ak) are excluded; constant omitted. N is smaller than 179 because some subjects did not report major and gender. Major in both science and economics is used here because there is only one subject majoring in economics in the sample.

which can also be seen in column (4), where neither CRT nor *Ak*×CRT has a significant effect on *Lk*. One possible reason could be that CRT might not be a good predictor of belief formation process in level-k thinking: recall that subjects need to think about strategies of other human subjects and also need to have the ability to calculate the iterative best responses. When the ability part in *Lk* is controlled by *Ak*, CRT on its own might not have a significant and high correlation with belief part in Lk, and thus CRT might turn out to be insignificant in the specifications (3) and (4).

Result 2.3. *More cognitively able subjects have higher ability bounds but not rationality bounds. Lkb subjects have higher cognitive ability than Lka subjects.*

2.5.3 Discussion

The reasoning ability of the online Mturk sample is significantly lower than that of the university student sample from Chapter 1 (mean Ak of 0.42 *vs*. 1.02, Mann–Whitney test $p = 0.000$). This could be explained by three possible reasons. First, the Mturk sample has a much lower cognitive ability, compared to the student sample.[10] Given the positive relationship between CRT and ability bounds, it might not be a surprise that the Mturk sample performed much worse than the student sample. Second, the Mturk sample is likely to exert lower effort resulting in noisier data (see Hauser et al., 2019; Aguinis et al., 2021, for surveys). Therefore, it might be reasonable to screen out subjects who spent too little time on the games, which is used as a proxy for effort. However, even after the screening, Mturk sample still performed worse (see Table B.3 and B.4). Third, the monetary incentive given to the Mturk subjects was lower, compared to that given to the subjects in the lab. It has been shown that a higher stake could induce higher reasoning depth in the lab (Alaoui and Penta, 2016; Alaoui et al., 2020). It is of interest for future research to study whether the small stake is the cause of the worse performance of the Mturk sample.

On the other hand, it is puzzling that the behavior of the student sample and of the Mturk sample is similar in repeated public goods and prisoner's dilemma games,

[10] The Mturk sample has a mean correct rate of CRT 0.39, significantly lower than 0.65 of the student sample (Mann–Whitney test $p = 0.000$). Note that the correct rate is used here for comparison instead of the score because Chapter 1 used 4 CRT questions based on Toplak et al. (2011) while this chapter used 3 questions based on Frederick (2005), three questions based on Toplak et al. (2011) and one additional question (as in Duffy et al. (2022)).

dictator game and trust game (e.g., Paolacci et al., 2010; Horton et al., 2011; Amir et al., 2012; Berinsky et al., 2012; Goodman et al., 2013; Arechar et al., 2018; Snowberg and Yariv, 2021; Duffy et al., 2022), but not in beauty contest games (Capra, 2019) or ring games. One possible explanation could be that level-k thinking involved in the two latter two types of games is strategically more complex than the reasoning in the other games mentioned above. Therefore, under what conditions one can replicate lab experiments on level-k thinking using the online sample might be another interesting avenue for future research.

2.6 Conclusion

This paper studies whether Lk behavior is driven by limited reasoning ability or low order of belief on opponents' rationality. In the within-subject experiment conducted on Mturk, subjects played a set of ring games facing two different types of opponents simultaneously - other Mturk subjects and themselves. The design allows one to measure each subject's reasoning steps with and without strategic uncertainty, respectively. Importantly, the payoff tables are designed to provide subjects incentive to reason until hitting their ability bounds.

I find that 98% of the subjects cannot play Nash equilibrium and have bounded reasoning ability. Further, 76% of the subjects are found to be ability-bounded Lk players while 10% of the subjects are belief-bounded Lk players. I also find that subjects who obtain higher CRT scores tend to have higher ability bounds but not rationality bounds, and are more likely to be Lkb. To conclude, this study documents evidence that subjects are largely bounded by their reasoning ability, which is the primary driver of low rationality bounds, as opposed to non-equilibrium belief.

Chapter 3

Inexact Information, Strategic Sophistication and Equilibrium Transition: A Quasi-Continuous-Time Experiment

with Zhi Li and Jianxun Lyu

3.1 Introduction

Understanding whether and how agents can move out of bad equilibria to more efficient ones is critical for the improvement of social welfare. Recent advancements in the evolutionary game theory show that sufficiently "inexact" (i.e., inaccurate but unbiased) information about opponents' behaviors can facilitate spontaneous transitions among strict Nash equilibiria (Ellison, 1997; Sandholm, 2001; Oyama et al., 2015; Sawa and Wu, 2023) (hereafter the inexact information models). The intuition is that agents might form false beliefs that the majority of a population has deviated from an established equilibrium even if only a small proportion has done so, and hence voluntarily deviate from a strict equilibrium. However, the transitions in these models rely on two key assumptions: first, agents best respond to signals regardless of the accuracy of the signals; and second, agents are myopic so that they only care about their immediate but not long-term gains or losses. It is critical to know whether human subjects follow such assumptions and whether the inexact information can facilitate efficiency-improving transitions among strict Nash equilibria empirically. Answering these questions is very much relevant in the era of social media where information is easier to be transmitted and manipulated.

This paper looks into the role of information about opponents' behaviors in equilibrium transitions in a quasi-continuous-time lab experiment. We investigate whether it is more or less accurate information that facilitates equilibrium transitions and how agents manage to do so if a transition succeeds. We design the experiment and derive theoretical predictions based on sampling best response dynamics (sBRD) (Oyama et al., 2015). In sBRD, a population of agents play a coordination game repeatedly, observe a private random sample of the population, and then play a myopic best response to their private samples. In our experiment, 14 subjects played a three-action coordination game with three strict equilibria differing in efficiency, with the initial state being the least efficient equilibrium. We control the information by varying the sampling size s, that is, $s = 2$ *vs.* $s = 7$. The theoretical prediction of sBRD is that the population (group) transits away from the least efficient equilibrium to the most efficient one under the treatment $s = 2$. In contrast, under the treatment $s = 7$, the population is expected to transit to the medium-efficient equilibrium or stay at the status quo, i.e., the least efficient equilibrium.

This paper finds novel patterns of equilibrium transitions with inexact information. All the session groups are observed to transit away from the least efficient equilibrium. However, in contrast to the sBRD predictions, transitions to the most efficient equilibrium are more frequent under $s = 7$ than $s = 2$. Specifically, all of the session groups under $s = 7$ transited to the most efficient equilibrium, while only one half under $s = 2$ transited to the most efficient one with the other half to the medium-efficient equilibrium.

The reason for the inconsistency between our experimental results and the theoretical predictions is twofold. First, different from the theory assumption, we find that subjects' responses do depend on the accuracy of the information: subjects tend to be less responsive to changes in signals with a smaller sampling size. In both treatments in our experiment, the probability of playing an action increases with the fraction of that action that subjects observe in their signals, but subjects under $s = 7$ adjust their actions more promptly than those under $s = 2$. The slower responsiveness associated with inaccurate signals results in more deviations from myopic best responses that point to a less efficient equilibrium (downward deviations hereafter) under $s = 2$, which delay or prevent transitions from the medium-efficient equilibrium to the most efficient one. Second, we find that more accurate information induces more long-term strategic behaviors, i.e., persistent deviations in the direction towards a more efficient equilibrium (upward deviations hereafter), resulting in more successful transitions to the efficient Nash equilibrium under $s = 7$ than $s = 2$.

The long-term strategic behaviors of persistent upward deviations induced by more accurate information in our results support the models of strategic teaching (Camerer et al., 2002; Naidu et al., 2010; Lyu, 2022). In these models, forward-looking agents teach myopic ones to play the efficient equilibrium by persistent upward deviations from an inefficient equilibrium. Lyu (2022) provides a theoretical model that sufficiently accurate information is necessary for strategic upward deviations, since forward-looking agents are willing to do so only if the opponents can observe such deviations and if forward-looking agents can observe whether the opponents indeed follow to deviate in the future. As a further evidence, we find that upward deviations are less correlated with the expected immediate payoff loss than downward deviations, indicating that farsighted subjects think beyond the immediate loss of upward deviations.[1]

Our results on the correlation between deviations and the payoff loss of deviations also contribute to the experimental literature on noisy suboptimal choices in games.

[1]This finding also implies that the equilibrium transitions observed in our experiment are different from those arising from random mistakes as in the models of stochastic stability (Foster and Young, 1990; Young, 1993; Kandori et al., 1993).

For example, Battalio et al. (2001); Bilancini et al. (2020) find that deviations from myopic best responses on average are cost-dependent, but Lim and Neary (2016) find that deviations are not cost-dependent. The mixed experimental evidence might be explained by the differentiation between upward and downward deviations found in our experiment. Downward deviations arise from slow responsiveness to the inaccurate signals, and thus are strongly affected by the cost of deviations; while upward deviations are more likely to be intentional and hence are less sensitive to the cost.

This paper has a methodological contribution to the design of evolutionary game experiments. The initial population state plays a key role in experiments of equilibrium transitions (or evolutionary games in general). In order to place all the subjects in a certain state, a common approach in the existing literature is to induce subjects to hold private preference to play certain action (e.g., Smerdon et al., 2019; Duffy and Lafky, 2021; Andreoni et al., 2021). In contrast with this approach, we propose a simple protocol of assigning subjects a default action in the beginning of the game. We find this protocol effective in a quasi-continuous-time experiments where subjects make choices asynchronously, and thus a population is less likely to jump away from the assigned initial state.

This paper also relates to a growing literature on experimental studies of social changes. Smerdon et al. (2019), Duffy and Lafky (2021), and Andreoni et al. (2021) study the social changes with perfect and imperfect information on individuals' private preferences over two social states in discrete-time lab experiments. More information about opponents' preferences is found to be beneficial for efficiency-improving social changes, since imperfect information might hinder social changes by invoking pluralistic ignorance - individuals are not sure whether the majority of a society prefers the status quo or the alternative state. We find inexact information about opponents' behaviors generates similar effects in our experiment when subjects have common knowledge about preferences.

It is worth noting the difference between our research and the experimental literature on the role of information in equilibrium selection based on global games (Heinemann et al., 2004; Anctil et al., 2004; Cornand, 2006; Cabrales et al., 2007; Van Huyck et al., 2018). The latter focuses on testing the prediction of static equilibrium plays, where subjects receive exogenously pre-specified imperfect (private or public) signals about a payoff-relevant state. We investigate transition dynamics where subjects are uncertain about the opponents' current behaviors and the received signals depend on the choices of their opponents.

The rest of the paper is organized as follows. Section 3.2 presents the experimental design, hypotheses, and the implementation. In section 3.3, we discuss the main findings of the experiment. The last section concludes.

3.2 Experimental Design and Hypotheses

3.2.1 The Game

A population of agents are randomly matched to a two-person coordination game recurrently. The coordination game has an action set $S = \{1, 2, 3\}$ and the payoff matrix is given by

$$G = \begin{pmatrix} 20 & 12 & 6 \\ 12 & 24 & 18 \\ 0 & 14 & 30 \end{pmatrix},$$

where G_{jk} corresponds to the payoff of action $j \in S$ when the opponent plays action $k \in S$. $X = \{x \in \mathbb{R}_+^{|S|} : \sum_k x_k = 1\}$, the simplex in $\mathbb{R}^{|S|}$, is the set of population states (or mixed strategies); for each population state $x \in X$, x_k is the fraction of players

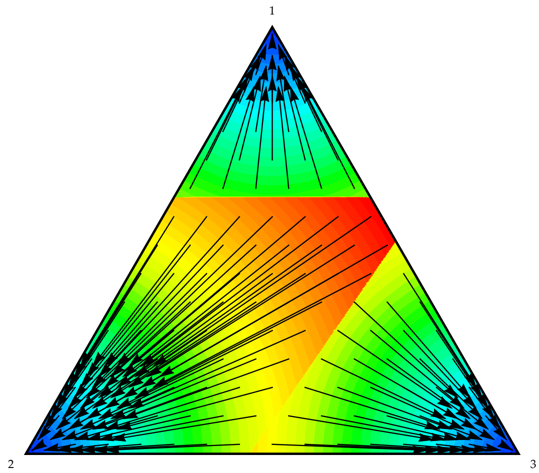
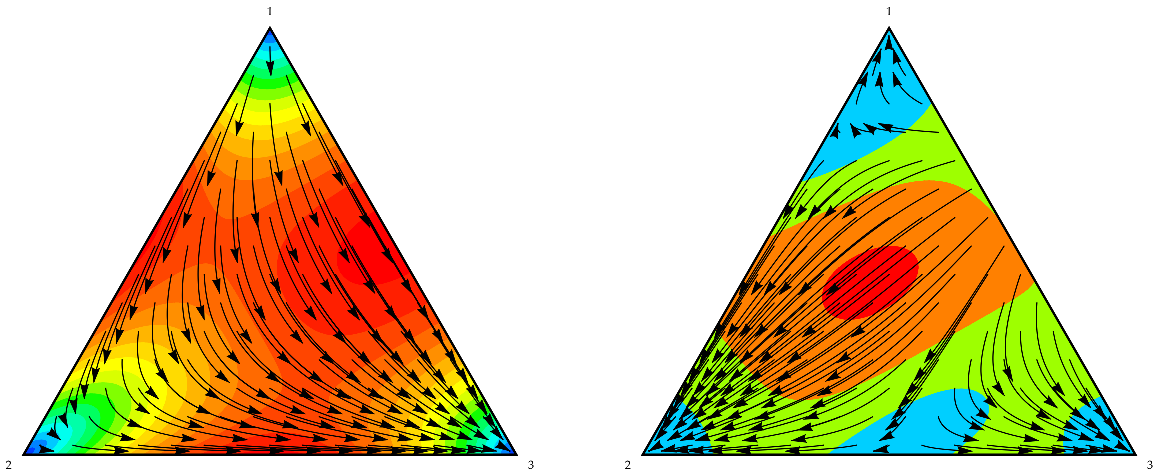


Figure. 3.1: Phase diagram of BRD.



A. $s = 2$ **B**. $s = 7$

Figure. 3.2: Phase diagrams of sampling BRD.

choosing action $k \in S$; e_k for each $k \in S$ represents the population state where all players use action k. In the random matching environment, the expected payoffs of actions are given by a function $F(x) = Gx \in \mathbb{R}^{|S|}$, where the kth element of $F(x)$ is the expected payoff to action k at population state x.

The game has three strict Nash equilibria of e_1, e_2, and e_3, ranked by efficiency in an ascending order. The question is whether and how a population can transit from e_1 to a more efficient state. If players have the full information about the population state x and play a myopic best response to x, the evolution of population state can be modeled by the standard Best Response Dynamic (BRD) (Gilboa and Matsui, 1991). The three monomorphic states e_1, e_2, e_3 are locally stable under BRD (Figure 3.1).[2] Thus, a transition away from e_1 is unlikely to happen with the full information if all players are myopic.

Note that the game used in the experiment has no fully mixed Nash equilibrium. Therefore, the best response regions of the three actions (or the basins of attraction of strict Nash equilibria) only intersect one by one, as shown in Figure 3.1. It suggests that to transit away from e_1 to e_3, a population must first transit to e_2. In addition, it is worth noting that along the edge e_1e_2, the tipping point of transitions from e_1 to e_2, T_{12}, is 0.4, i.e., it requires more than 40% of the population to play action 2 to transit from e_1 to e_2. Similarly, along the edge e_2e_3, the tipping point is roughly 0.45. Therefore, the transition from e_2 to e_3 is more difficult than that from e_1 to e_2.

When agents only observe a finite sample of the population, and thus play a myopic best response (mBR) to the empirical distribution of actions in the sample, the population might be able to transit to a more efficient equilibrium under the sampling Best Response Dynamics (Oyama et al., 2015). Panels A and B in Figure 3.2 illustrate the sampling dynamics of the game with a sampling size of $s = 2$ and $s = 7$, respectively. When $s = 2$, all of the interior trajectories converge to e_3. When $s = 7$, the stability of the dynamics is similar to the full-information case, but the basin of attraction of e_1 becomes smaller, suggesting an easier transition from e_1 to e_2 than with the full information.[3]

The theoretical predictions above cannot be directly applied to a lab environment where the population size is relatively small. The reason is that sampling BRD can be seen as a deterministic approximation of a stochastic process when the population size is sufficiently large. When the population size is small, e.g., small groups in lab

[2]Phase diagrams in this paper are plotted in Dynamo developed by Franchetti and Sandholm (2013).

[3]In fact, for the game used in this paper, all strict Nash equilibria are locally stable for $s > 2$, and e_3 is almost globally stable for $s = 2$. The choice of the two sampling sizes 2 and 7 is further determined by the differences in the transition probabilities under the two sampling sizes (see Table 3.1 and 3.2).

Table 3.1: Transition probabilities from e_1 to e_2.

# of Deviations	1	2	3	4	5
$s = 2$	0.838	0.971	0.992	0.999	1.000
$s = 7$	0.016	0.192	0.614	0.902	0.989

Notes: $T_{12} = 0.4$, $N = 14$, one-shot deviation from action 1 to 2.

Table 3.2: Transition probabilities from e_2 to e_3.

# of Deviations	1	2	3	4	5
$s = 2$	0.823	0.961	0.995	1.000	1.000
$s = 7$	0.000	0.000	0.014	0.105	0.363

Notes: $T_{23} = 0.45$, $N = 14$, one-shot deviation from action 2 to 3.

experiments, there is non-negligible stochastic components that might affect transition probabilities. We estimated transition probabilities in small populations under different sampling sizes. Tables 3.1 and 3.2 respectively present the transition probabilities in 1000 simulations from e_1 to e_2 and from e_2 to e_3 for the population size $N = 14$ given the number of one-shot deviations. The estimations are generally consistent the predictions of sampling BRD with one exception that it is likely to transit from e_1 to e_2 with $s = 7$. However, the population is still unlikely to transit from e_2 to e_3 under $s = 7$ unless there are sufficiently many deviations.

3.2.2 Testable Hypotheses

We derive the testable hypotheses from the theoretical predictions of sampling BRD with e_1 as the initial state. Hypothesis 3.1 describes the equilibrium transitions resulting from transition probabilities under the two sampling sizes presented in Tables 3.1 and 3.2 (given a small number of deviations).

Hypothesis 3.1. *Transitions among the three strict Nash equilibria:*

(a) The population will transit away from e_1 more frequently under $s = 2$ than under $s = 7$.

(b) The population will transit to e_3 more frequently under $s = 2$ than under $s = 7$.

Hypothesis 3.2 tests whether subjects play myopic best response given the inexact information regardless of the accuracy of the information determined by the sampling size.

Hypothesis 3.2. *Subjects best respond to the empirical distribution of actions in their private samples regardless of the sampling size.*

Hypothesis 3.2 has an important implication for equilibrium transition. If the rate of best response depends on the accuracy of information, equilibrium transition predicted by sBRD might fail to appear. The reason is that the information indicating an action corresponding to a more efficient equilibrium as the mBR could be rare at the early stage of transitions. If subjects do not best respond to such information immediately, it will require more persistent efficiency-improving deviations from mBR to initiate a transition. As those deviations are costly, myopic players might not be willing to deviate, in which case, an efficiency-improving transition might fail.

3.2.3 Experimental Design and Implementation

In each session, a group of 14 subjects played the game as described above for 80 periods with a sampling size of either $s = 2$ or $s = 7$. Each period lasted for 10 seconds. At the beginning of a period, every subject received a private signal, which consists of s random draws of their opponents' action choices in the last period with replacement (see Figure C.10 in Appendix C.4 for the interface). Hereafter, we denote a signal by $\hat{x} = (s_1/s, s_2/s, s_3/s)$, where s_k is the sampled number of action k in the signal with $\sum s_k = s$.[4] Upon receiving the signals, subjects decided whether to choose an action or not in the current period. The decisions in the last period were carried over if no

[4]Here $\hat{x}$ is the empirical distribution of actions in the received random sample. It is also an unbiased estimator of population state x.

decision was made. That is, subjects were free to change their previous decisions in a period in the sense that the decision-making in our design was in the quasi-continuous environment and thus asynchronous. The population state and payoff were updated at the end of each period. Importantly, neither the feedback about subjects' own payoff nor that about their opponents' was provided during the game to avoid potential confounding effects - the payoff feedback might be perceived as the complementary information for decision-making.

To set e_1 as the status quo, we employed a novel design of default actions - action 1 was assigned to every subject as the default option before the game started.[5] Note that the signals in the first period were based on the default state e_1, and thus all the subjects received a signal $(1, 0, 0)$, which indicates that action 1 is the myopic best response. As a result, myopic subjects would always play action 1 and thus be stuck at e_1 if there is no disturbance such as mistakes or intentional deviations from mBR. Further, the asynchronous move combined with the 10-second time limit makes abrupt large deviations from the default state practically difficult in the first period, so the state quo is expected to be sustained. Thus, the default option design can ensure e_1 as the state quo.[6]

The accumulated payoff of all the 80 periods was paid. In the end of the experiment, subjects answered a paid survey, including self-reported demographic information, cognitive reflection tests (Frederick, 2005; Toplak et al., 2011), social preferences (Falk et al., 2016, 2018), incentivised risk preferences (Eckel and Grossman, 2002; Reynaud and Couture, 2012) and level-k thinking using the 11-20 game (Arad and Rubinstein, 2012).

The experiment was conducted at the Finance and Economics Experimental Laboratory at Xiamen University in China using oTree (Chen et al., 2016). A total of 112 subjects in 8 sessions were recruited from Xiamen University. Each subject only

[5] Subjects were informed that they would be assigned an action and only knew which action was assigned to themselves right before the game started.

[6] Our results show that by using a default option every session stayed inside of basin of attraction of e_1 in the first period.

participated in one session, and each session lasted for about 45 minutes. The instructions were read aloud in front of the subjects.[7] Subjects had to pass a comprehension quiz and had an unpaid 10-period training session before proceeding to the game with payments. The average earnings were 26 CNY ($\approx$ 4.02 USD) per subject, close to a double of students' hourly wage rate.[8] For summary statistics of our data set, see Appendix C.2 Table C.2.

3.3 Results

3.3.1 Transitions

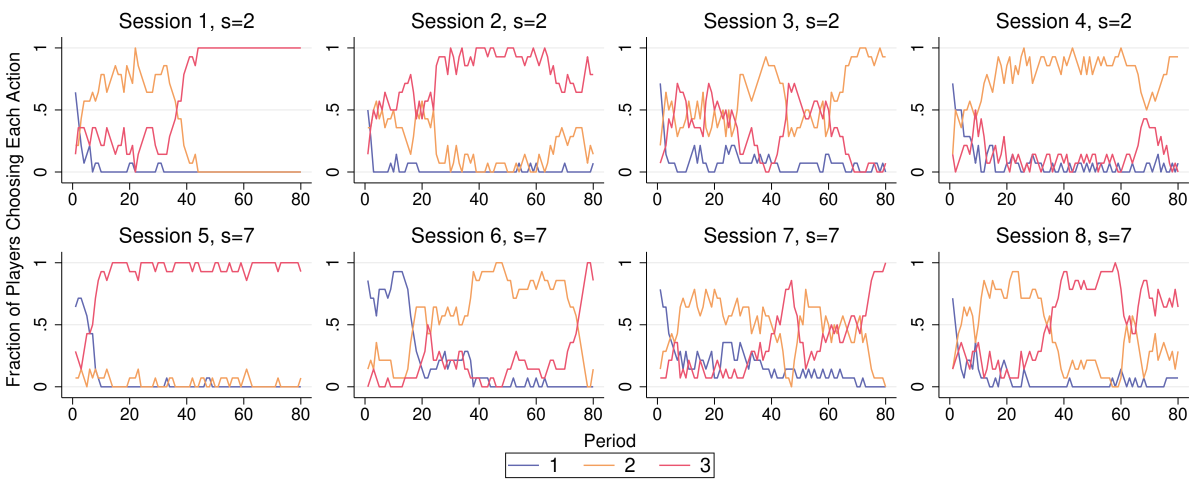
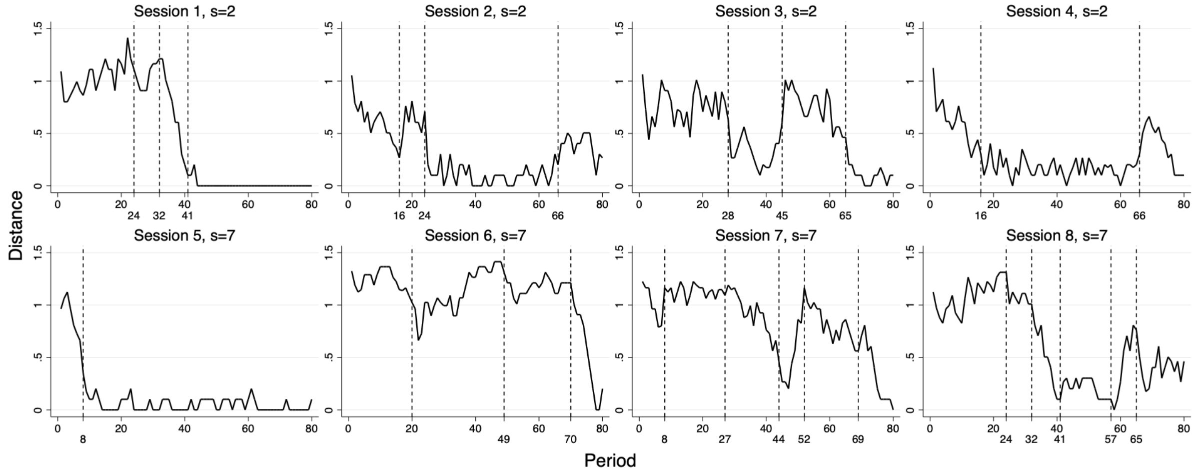


Figure. 3.3: Evolution of fractions of actions by sessions.

Figure 3.3 plots the evolution of the fractions of players choosing each action in each session, and it provides visual evidence on the patterns of equilibrium transitions. The fraction of action 3 (the red curve) is close to 1 at the end of the game in all of the sessions under $s = 7$; while under $s = 2$, only two groups (Sessions 1 and 2) transited to

[7] A copy of instructions translated in English is in Appendix C.3.

[8] The hourly students' wage rate at Xiamen University was 18 CNY at the time of the experiment.



Notes: The vertical reference lines are the estimated structural breaks.

Figure. 3.4: Distance of the population state to the equilibrium each session arrived at the end of the session.

e_3 with the other two (Sessions 3 and 4) to e_2.[9] In particular, there exists one session under $s = 7$ (Session 5) that the group transited away from e_1 to e_3 directly and rapidly without passing through the neighborhood of the medium-efficient equilibrium e_2.

Result 3.1. *Both inexact information treatments induce transitions away from the least efficient equilibrium to a more efficient equilibrium; the transitions to the most efficient equilibrium e_3 are more frequent with more accurate information of $s = 7$ compared to $s = 2$.*

To see if a group indeed *converges* to or has a tendency to converge to the equilibrium reached at the end of the game, we look at the distance of the population state $x = (x_1, x_2, x_3)$ to the final equilibrium reached e_2 for sessions 3 and 4, and e_3 for the other sessions. A significant decreasing trend in distance, combined with the final arrival at an equilibrium (or neighborhood of an equilibrium) can be interpreted as a *convergence* to the equilibrium arrived. As shown in Figure 3.4, in each session, the distance has apparent abrupt changes at some points. Thus, we estimate structural breaks for each session

[9]Here transitions to an equilibrium include cases where populations transit to a small neighborhood of the equilibrium.

Table 3.3: OLS of distance on time allowing for breaks in trend and constant.

Distance	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t1	0.015***	-0.035***	0.001	-0.039***	-0.091***	-0.006	-0.040**	0.017***
	(0.002)	(0.005)	(0.003)	(0.005)	(0.010)	(0.004)	(0.015)	(0.003)
t2	0.044***	0.003	0.000	-0.001	-0.001	0.022***	-0.000	-0.004
	(0.011)	(0.015)	(0.007)	(0.001)	(0.000)	(0.002)	(0.004)	(0.015)
t3	-0.138***	-0.000	-0.024***	-0.046***		0.005	-0.038***	-0.099***
	(0.009)	(0.001)	(0.005)	(0.006)		(0.004)	(0.005)	(0.013)
t4	-0.001	-0.014**	-0.005			-0.122***	0.133***	-0.013**
	(0.001)	(0.007)	(0.008)			(0.011)	(0.015)	(0.005)
t5							-0.025***	0.118***
							(0.005)	(0.015)
t6							-0.082***	0.006
							(0.009)	(0.006)
constant1	0.854***	0.912***	0.749***	0.927***	1.243***	1.297***	1.210***	0.881***
	(0.030)	(0.052)	(0.053)	(0.049)	(0.052)	(0.045)	(0.077)	(0.041)
constant2	-0.194	0.579*	0.315	0.197***	0.074***	0.357***	1.136***	1.169***
	(0.316)	(0.314)	(0.253)	(0.040)	(0.019)	(0.078)	(0.077)	(0.429)
constant3	5.748***	0.117*	2.063***	3.769***		0.901***	2.263***	4.135***
	(0.343)	(0.059)	(0.296)	(0.459)		(0.211)	(0.177)	(0.465)
constant4	0.076	1.391***	0.430			9.750***	-5.885***	0.835***
	(0.063)	(0.481)	(0.598)			(0.808)	(0.736)	(0.262)
constant5							2.325***	-6.780***
							(0.298)	(0.922)
constant6							6.514***	-0.069
							(0.703)	(0.424)
N	80	80	80	80	80	80	80	80

Notes: Columns (1)-(8) are regressions of eight sessions respectively. Independent variables t1-t6 and constant1-constant6 are separated time trends and constants based on estimated structural breaks. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(Ditzen et al., 2021), which allow us to estimate trends at different stages.[10] Table 3.3 columns (1) to (8) report OLS regressions of distance on time with the estimated breaks respectively for the eight sessions and all support the convergence to the equilibrium reached toward the end of each session. For example, Session 6 is estimated to have three breaks in periods 20, 49 and 70 - the distance of population state to e_3 is large and

[10]Both breaks in trend and constant are allowed in the test and estimation of structural breaks in distance series.

relatively stable before period 70 but it has a significant decreasing trend since period 70, indicating a convergence to the most efficient equilibrium e_3 in the end (in Table 3.3 column (6), the time trend after period 70, i.e., t_4 has a significantly negative coefficient, indicating that the distance of population state to the final equilibrium identified is decreasing over time).[11]

Result 3.1 does not support Hypothesis 3.1 which predicts more transitions to e_3 under a smaller sampling size of $s = 2$. Hence it raises two questions that sampling best response dynamics fail to explain: first, why all of the groups transited to e_3 under $s = 7$, contradicting the prediction of none transiting to e_3 under $s = 7$; and second, why some of the groups failed to transit from e_2 to e_3 within 80 periods under $s = 2$, in contrast to the prediction of all transiting to e_3 under $s = 2$. To answer these two questions, we focus on investigating the differences in individual behaviors across the treatments, specifically the deviations from mBR to signals in the remainder of the section.

3.3.2 Deviations from mBR to Signals

We define an action choice of subject i in period t that is not mBR to the private signal as a *deviation*. We find that there is a significant portion of deviation behaviors at the beginning of the game under both treatments, and the deviation rate has a decreasing trend over time (Figure C.2).[12] However, we find no significant difference in the aggregate deviation rates across the treatments according to two-sided Mann-Whitney test (0.229 vs.

[11] Note that the reversion in Session 8 in the end, where the group slightly moves away from the equilibrium reached, is not significant. Both trend and constant in the final segment of the session (after the last break) are not significant (column (8) of Table 3.3), indicating the distance to e_3 in Session 8 is not significantly different from zero, and thus the reversion is not significant. Since e_3 Pareto dominates the other two equilibria, it is expected that no group would deviate from e_3 once it is reached, and the result might be more robust in sessions with more periods.

[12] The figure points out the importance of the deviations for the transitions - the deviation rate rises when the transition starts, and falls when a Nash equilibrium is reached. It also explains the reversion in the aggregate deviation rate at the treatment level in the last 20 periods, as in some sessions the transitions happened at the end of the game.

0.240, $p = 0.4472$). This observation motivates us to look at the directions of deviations, the differences in which might favor different equilibrium transitions.

3.3.2.1 Direction of Deviations

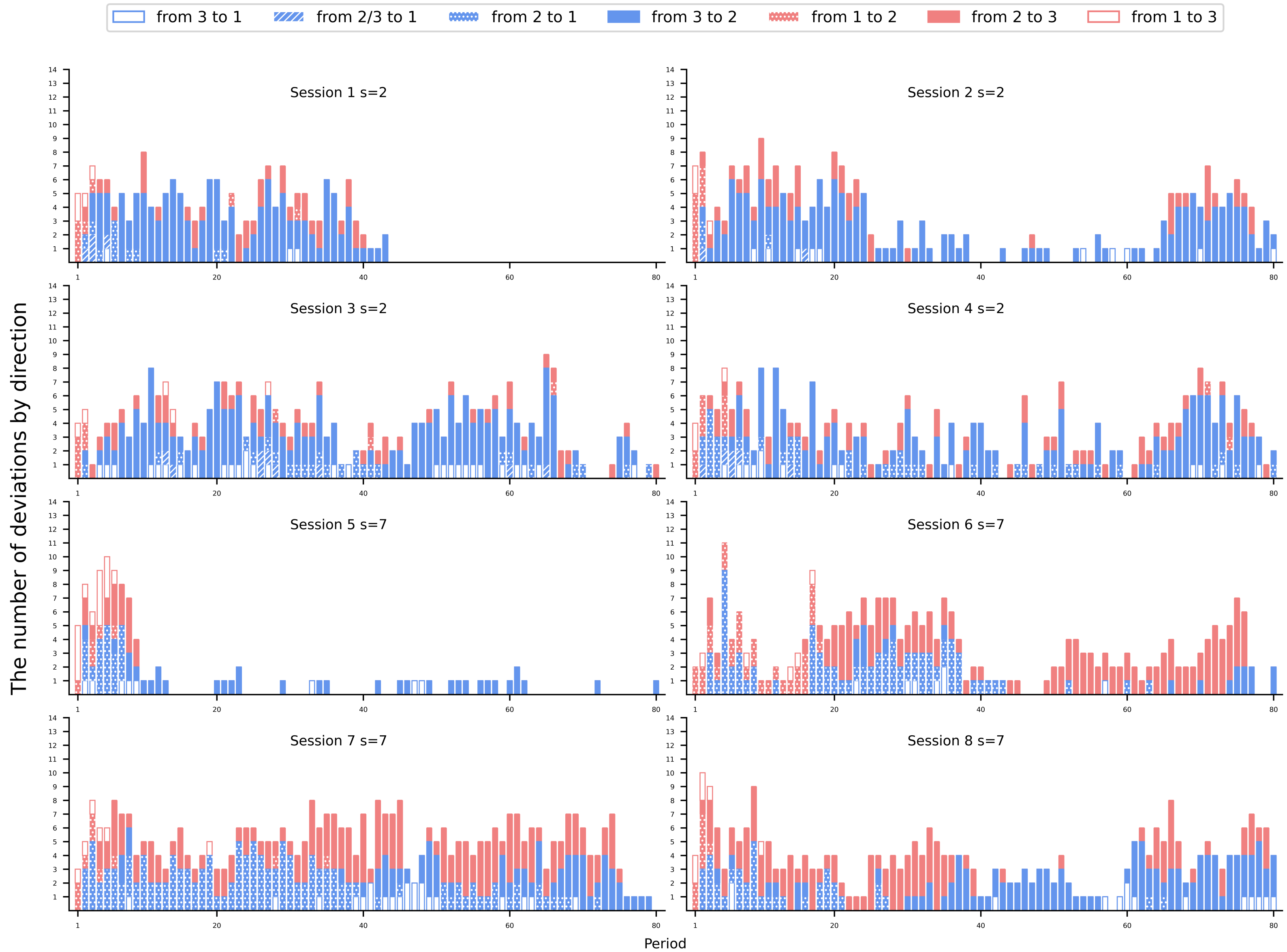
As pointed out in section 3.2.1, the deviations *towards a more efficient equilibrium* are of significant importance for efficiency-improving transitions. In contrast, deviations *toward a less efficient equilibrium* would hamper such transitions. Consider an example where the signals indicate action 2 as the mBR. In this example, deviations to action 3 obviously favor transitions to e_3 while deviations to 1 do not. That is, the direction of deviations matters. Formally, we define the direction of deviations as follows.

Definition 1. *Given a private signal $\hat{x}$, an action $a \in S$ is called an upward (downward) deviation from the myopic best response $a^*(\hat{x}) \in S$ if $a > a^*(\hat{x})$ ($a < a^*(\hat{x})$).*

Figure 3.5 illustrates the number of deviations by direction in each session over time, where upward (downward) deviations are highlighted in red (blue). The first pattern to notice is that the treatment $s = 7$ has fewer downward deviations (51.81% vs. 76.41%) and more upward deviations (48.19% vs. 23.59%) out of all the deviations than the treatment $s = 2$ (χ^2 test $p < 0.001$, also see Table 3.4).[13] In particular, the treatment $s = 2$ has more downward deviations from action 3 to 2 (60.72% vs. 21.73%, χ^2 test $p < 0.001$) and fewer upward deviations from action 2 to action 3 (18.91% vs. 40.48%, χ^2 test $p < 0.001$) than the treatment $s = 7$. This is consistent with the transition result that there are more transitions to e_3 under $s = 7$ than $s = 2$.

In addition, focusing on upward deviations, the deviations $1 \to 2$ are more frequent at the early stage while $2 \to 3$ are more frequent at the later stage of the game. It is consistent

[13] If taking best responses into account, in treatment $s = 2$ there are 77.10% of best responses, 17.50% of downward deviations, and 5.40% of upward deviations. While the proportions are 75.96%, 12.46% and 11.58% respectively in treatment $s = 7$. Still the difference is significant at the 1% level according to χ^2 test ($p < 0.001$).



Notes: Upward (downward) deviations are in red (blue). Under treatment $s = 2$, there exists one signal $(1/2, 0, 1/2)$ admitting both action 2 and 3 as the best response. Deviations conditional on this signal are labelled by the legend 'from 2/3 to 1'.

Figure. 3.5: Direction of deviations over time by session.

Table 3.4: Percent of each type of deviations by treatment.

		$s = 2$	$s = 7$
Down	from 2/3 to 1	2.05	-
	from 3 to 1	5.46	4.74
	from 3 to 2	60.72	21.73
	from 2 to 1	8.19	25.35
	Subtotal	76.41	51.81
Up	from 1 to 2	3.22	4.83
	from 2 to 3	18.91	40.48
	from 1 to 3	1.46	2.88
	Subtotal	23.59	48.19
	Total	100.00	100.00

with the observation that populations often first transit from e_1 to a neighborhood of e_2, then to e_3, as shown in Figure 3.3. This also explains the hump-shape distribution of deviations over time, as subjects attempted to transit to the most efficient equilibrium e_3 in two steps (first e_1 to e_2 then e_2 to e_3, except Session 5). Such deviation pattern can be explained by the negative correlation of a deviation play and its associated payoff loss, which will be discussed in section 3.3.5 in detail: For example, at the early stage of the game when more subjects play action 1, a deviations $1 \to 3$ in fact results in a higher payoff loss than a deviations $1 \to 2$, therefore, it is reasonable to observe more upward deviations from 1 to 2 at the early stage.

Result 3.2. *There are significantly more upward deviations and fewer downward deviations under a larger sampling size $s = 7$ than a smaller one $s = 2$.*

3.3.2.2 Treatment Effects on Deviation Frequency by Direction

One possible explanation for the significant treatment difference in the quantities of upward and downward deviations is that there exists a significant difference in the frequencies of upward and downward deviation plays across treatments. Here we examine this possibility.

Let us first denote $Down_i$ and Up_i as the frequency of downward and upward deviations of subject i:

$$Down_i = \frac{D_i}{TD_i}, \qquad Up_i = \frac{U_i}{TU_i}.$$

where D_i and U_i denote the number of periods subject i played downward and upward deviations, respectively; TD_i and TU_i denote the total periods of downward and upward deviations available to i, respectively. Note that TD_i and TU_i are in general less than 80. For example, if a subject i observes a signal in period t that indicates action 1 as mBR, then a downward deviation is not an option for subject i in period t.[14]

Figure 3.6 shows the empirical cumulative distributions of Up_i and $Down_i$ by treatment. On average, subjects under $s = 2$ seem to play downward deviations more often than subjects under $s = 7$, while the opposite holds for upward deviations.[15] Both Mann-Whitney test ($p = 0.005$) and Kolmogorov-Smirnov test ($p = 0.006$) suggest the significant difference in $Down_i$ across treatments, but not in Up_i (Mann-Whitney test $p = 0.0684$, Kolmogorov–Smirnov test $p = 0.153$). It indicates that subjects do not play upward deviations more often under $s = 7$ than $s = 2$.

To further verify the findings, we run fractional logit regressions where the dependent variable is either $Down_i$ or Up_i. For both regressions, we use two specifications with model (1) including only $s7_i$, a dummy variable, taking the value of 1 if subject i was

[14]See Figure C.3 in Appendix C.2 for the chances that subjects in each session had to play downward and upward deviations. The chances to deviate downward has little difference across sessions, while the chances to deviate upward vary a lot, indicating the importance of the accommodation of available chances.

[15]Note there is one obvious outlier in the distribution of downward deviations for $s = 7$, which is equal to 0.86. With the outlier, the mean, sd, minimum, maximum of $Down_i$ are 0.13 0.15, 0, and 0.86, respectively. Without the outlier, those of $Down_i$ are 0.12, 0.12, 0, and 0.38, respectively. The outlier is omitted for the relevant regression analysis.

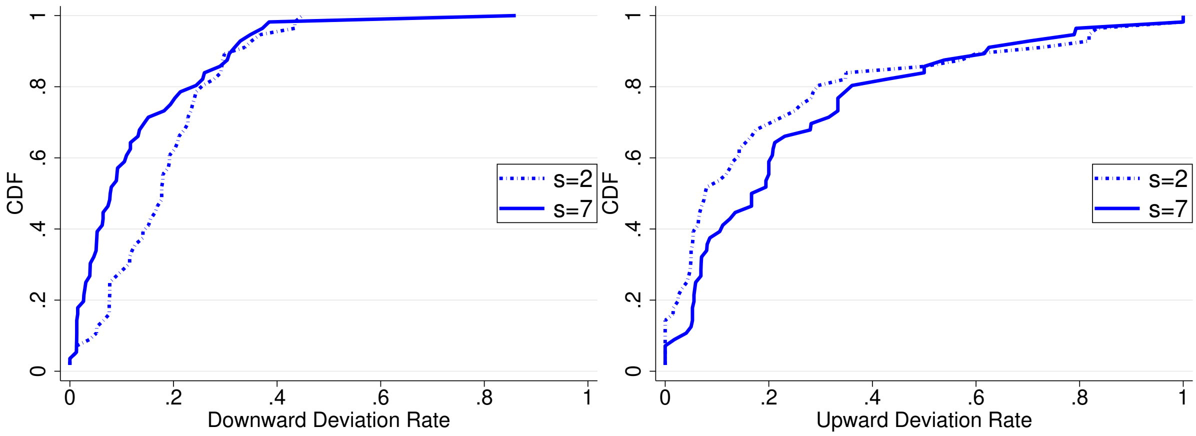


Figure. 3.6: Cumulative distributions of frequency of upward and downward deviations.

in the treatment $s = 7$ and 0 otherwise; and model (2) further controlling for personal characteristics including gender, major in economics, study level, age and cognitive ability, measured by the number of correctly answered seven CRT questions, normalized by the seconds spent on the questions scaled up by 100 (denoted by $nCRT_i$).[16] The standard errors are clustered at session level.[17]

The regression results are reported in Table 3.5. We find a significant negative effect of $s7_i$ on the frequency of downward deviations (columns (3) and (4) of Table 3.5), but no significant effect on the frequency of upward deviations (columns (1) and (2)), supporting the discussion above. Interestingly, high-cognition and male subjects tend to play upward deviations significantly more often and downward deviations less often. Additionally, subjects studying economics also deviate upward at a significantly higher frequency. The

[16]The normalized CRT (nCRT) instead of the raw CRT score is used due to few variations in the raw score with a mean of 5.50 and a standard deviation of 1.45. The normalized CRT has a mean of 2.41 and standard deviation of 1.40. The two measures show a reasonably high correlation at the 1% significance level, with a Pearson's correlation coefficient of 0.6167 and a Spearman's rho of 0.6981. See Figure C.4 for the relationship between the two measures in Appendix C.2. Using raw CRT score does not alter the results regarding the treatment effect.

[17]One could also use mixed logit models with a categorical dependent variable taking three different values for best responses, downward and upward deviations respectively. As shown in Table C.4 of the Appendix C.2, the results are qualitatively the same. It is important to note that one should not use multinomial logit due to the violation of the assumption of fixed choice set and thus IIA by design, since upward deviations and downward deviations are not always available, while mixed logit can deal with the issue.

Table 3.5: Test of treatment effect on deviation frequency by direction.

	Up_i		$Down_i$	
	(1)	(2)	(3)	(4)
$s7_i$	0.227	0.543	-0.481**	-0.577***
	(0.567)	(0.482)	(0.240)	(0.224)
$Female_i$		-0.598*		0.549**
		(0.327)		(0.215)
$Economics_i$		0.537***		0.073
		(0.156)		(0.278)
$nCRT_i$		0.150**		-0.141**
		(0.062)		(0.061)
$StudyLevel_i$		-0.691		0.178
		(0.457)		(0.311)
Age_i		0.108		-0.011
		(0.067)		(0.051)
Constant	-1.334**	-1.656**	-1.516***	-1.937**
	(0.528)	(0.835)	(0.130)	(0.770)
N	112	108	111	107

Notes: Fractional logit regressions. Standard errors are clustered at session level in parentheses. Fewer observations in columns (2) and (4) are due to missing values in $Female_i$ and Age_i. One outlier is omitted for columns (3) and (4) (see footnote 15). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

results are robust after controlling for the social preference and the level-k behavior (see columns (1) and (2) of Table C.3 in Appendix C.2).[18]

Result 3.3. *Subjects under the treatment $s = 7$ play downward deviations significantly less often than $s = 2$. However, the sampling sizes have no significant effect on the frequency of upward deviations.*

3.3.3 Forward-looking Players

In the section 3.3.2.1, we show that there are significantly more upward deviations in quantity under $s = 7$ but in the section 3.3.2.2, we find no significant difference in the

[18]We do not find an significant effect of $s7_i \times nCRT_i$, $s7_i \times Economics_i$, $nCRT_i \times Economics_i$ or $nCRT_i \times Female_i$ on Up_i or $Down_i$.

frequency of upward deviations across treatments. The expected number of upward deviations depends on both the probability of individual deviations and on the number of subjects who deviate. Since the probability (the frequency of upward deviations) does not differ significantly across treatments, a conjecture is that there are more subjects who are willing to play upward deviations under $s = 7$ than $s = 2$.

To test this argument, we first identify subjects who are forward-looking, namely who would like to play upward deviations intentionally in order to transit to a better equilibrium. To separate intentional deviations from mistakes, we focus on the persistent upward deviations (e.g., Lim and Neary, 2016). We define an upward deviation by subject i in period t to be strategic if she also plays an upward deviation in period $t + 1$.[19]

We classify subjects into three types based on their observed behaviors: myopic players, forward-looking players (henceforth FL players), and the others. The classification rule is summarized as follows (Table 3.6):

- **Myopic Players** play best responses for at least 85% of time with no strategic deviation, i.e., no persistent upward deviation;
- **FL Players** play best responses *and* upward deviations together for 85% of the time with at least one strategic deviation;
- **Others** who cannot be classified into the two types above.

The threshold of 85% is chosen according to the 95% confidence interval $(0.13, 0.17)$ of the mean time (0.15) of downward deviations in the 80 periods, which is set to allow for mistakes.[20]

[19]Note that there exist scenarios where one-shot upward deviation is not a mistake but cannot be identified as being strategic, because the choice of upward deviations might not be available in two consecutive periods given the signals. Therefore, our measure provides a lower bound of the number of strategic deviations.

[20]All of the downward deviations are considered as mistakes here. However, whether we consider down deviations as mistakes does not affect the identification of FL players who are defined by strategic deviations and best responses. In the next subsection, we will show that downward deviations are caused partly by the inaccuracy of signals.

Table 3.6: Classification of subject types. Count (Percent).

	FL Player	Myopic Player	Others
Classification Rule	BR + S + NS-Up $\geq$ 85% S $\geq$ 1	BR $\geq$ 85 % S = 0	None of the above
Session 1, $s = 2$, e_3	3 (21.43%)	6 (42.86%)	5 (35.71%)
Session 2, $s = 2$, e_3	4 (28.57%)	2 (14.29%)	8 (57.14%)
Session 3, $s = 2$, e_2	3 (21.43%)	1 (7.14%)	10 (71.43%)
Session 4, $s = 2$, e_2	4 (28.57%)	0 (0.00%)	10 (71.43%)
Total $s = 2$	14 (25.00%)	9 (16.07%)	33 (58.93%)
Session 5, $s = 7$, e_3	5 (35.71%)	8 (57.14%)	1 (7.14%)
Session 6, $s = 7$, e_3	9 (64.29%)	2 (14.29%)	3 (21.43%)
Session 7, $s = 7$, e_3	4 (28.57%)	2 (14.29%)	8 (57.14%)
Session 8, $s = 7$, e_3	7 (50.00%)	2 (14.29%)	5 (35.71%)
Total $s = 7$	25 (44.64%)	14 (25.00%)	17 (30.36%)

Notes: BR: best response; S: strategic deviation; NS-Up: nonstrategic upward deviation; NS-Down: nonstrategic downward deviation. The threshold 85% is chosen according to the 95% confidence interval (0.13, 0.17) of the mean downward deviation time. Each session is labeled by the treatment and the final equilibrium reached.

The requirement for the classification of FL players is not very strict in the sense that subjects will be classified as FL players if they reveal their type by only one strategic deviation. That is, it is not required that FL subjects play strategic deviations whenever possible. Moreover, it implicitly allows FL subjects to switch between upward deviations and best responses - one might stop deviating upward if she believes transitions would be unsuccessful. As demonstrated in Figure 3.6, few subjects deviate upward all the time. Further, we do not restrict subjects to upward deviate to the same action in two consecutive periods: subjects could deviate upward to action 2 and subsequently to action 3 if the transition to e_2 is perceived to be successful at some point.[21]

[21] One might argue that a smaller sample size would lead to smaller persistent chances for upward deviations, resulting in fewer subjects identified as FL players. In our data set, there are two such subjects who do not have two consecutive periods where they could deviate upward. One is identified as 'others' due to too many downward deviations while the other is identified as a myopic player with 7.5% of the time on one-shot upward deviations. The results do not change if the subject is classified as a FL player.

As shown in Table 3.6, 44.64% of the subjects are classified as FL players when $s = 7$, significantly greater than the proportion 25% when $n = 2$ (Fisher's exact test $p = 0.047$). Note that on average the percentage of FL players when $s = 7$ is very close to the tipping point for the transition from e_2 to e_3, namely 45% (t test $p = 0.9577$), which could explain the transitions to e_3 when $s = 7$.

Notice that a large number of subjects are identified as 'Others' under sampling size $s = 2$ compared to $s = 7$ (58.93% *vs.* 30.36%, Fisher's exact test $p = 0.004$). The subjects who are classified as 'Others' are due to a high downward deviation rate (26% on average). However, they still play mBR for most of the time in the experiment - the average rate of mBR of Others is 68.65%.[22] Thus, those subjects can be interpreted as 'conservative myopic players' because their downward deviations can be explained by their perception of inaccurate information and slow response to the signals. This will be explored in detail in the next section 3.3.4.

The treatment effect of sampling size on the probability of a subject being identified as a forward-looking player is estimated by the logit regressions, where the dependent variable is FL_i, a dummy taking the value of 1 if subjects i is identified as a FL player, and 0 otherwise. Table 3.7 presents two model specifications, with model (1) including only the treatment dummy of sampling size $s7_i$ and model (2) further controlling for individual characteristics of gender, major in economics, study level, age and normalized CRT score. The standard errors are clustered at the session level.

As shown in Table 3.7, there is a significant positive treatment effect of sampling size on the probability of a subject being identified as a forward-looking player under both specifications. This explains that the treatment $s = 7$ has more upward deviations and all transited to e_3 while there is no significant treatment difference in the frequency of upward deviations. Further, a subject with higher cognitive ability is also more likely

[22] Also see Figure C.7 in Appendix C.2 for the time spent on each kind of behavior across subject types at the individual level.

Table 3.7: Test of treatment effect on identified FL players.

FL_i	(1)	(2)
$s7_i$	0.884***	1.119**
	(0.314)	(0.464)
$Female_i$		-0.260
		(0.481)
$Economics_i$		-0.054
		(0.524)
$nCRT_i$		0.468***
		(0.131)
$StudyLevel_i$		0.110
		(0.589)
Age_i		0.020
		(0.128)
Constant	-1.099***	-2.983*
	(0.102)	(1.791)
N	112	108

Notes: Logit regressions. Standard errors clustered at session level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Fewer observations in column (2) due to missing values in $Female_i$ and Age_i.

to be a forward-looking player, but the effects of other individual characteristics such as gender and major that were found significant on the frequency of upward deviations are not significant on the probability of being FL players. Table C.3 in Appendix C.2 presents consistent results after controlling for patience, social preferences and level-k reasoning.[23]

Result 3.4. *A larger sampling size $s = 7$ induces significantly more subjects to play strategic deviations (or persistent upward deviations) than a smaller sampling size $s = 2$. Subjects with higher cognitive ability are more likely to be forward-looking players.*

Result 3.4 supports the conjecture based on strategic teaching models. Lyu (2022) suggests that if players have sufficiently accurate information about their opponents'

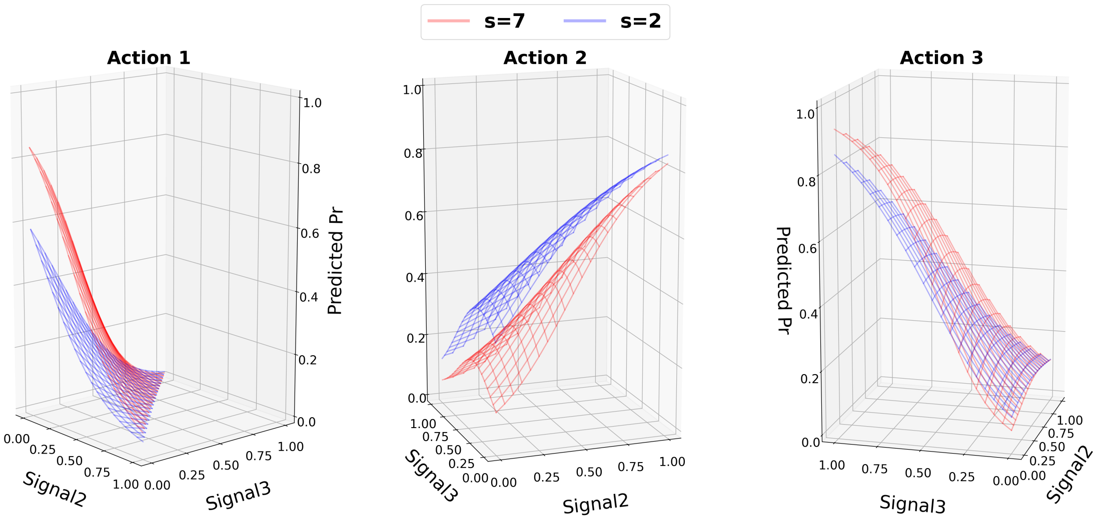
[23]We do not find an significant effect of $s7_i \times Female_i$, $s7_i \times nCRT_i$, $s7_i \times Economics_i$, $nCRT_i \times Economics_i$ or $nCRT_i \times Female_i$ on FL_i.

choices, forward-looking players with higher cognitive levels (measured by level-k reasoning in Lyu (2022)) will strategically deviate from an inefficient equilibrium at the early stage of a game to induce lower-cognitive-level players to deviate in the future so that the population can reach a more efficient equilibrium. This is because, with inaccurate information such as a sampling size 2, strategic deviations made by higher-cognitive-level players are less likely to be observed by lower-level players. Therefore, higher-level players will not play upward deviations in the first place in this case. The regression results in Table C.3, Appendix C.2 provide additional evidence for this argument with social preferences and level-k reasoning controlled. We find that the level-k reasoning also has a significant positive effect on upward deviations and on the probability of being a FL player.

Another channel through which the sampling size might affect strategic deviations is the existence of a free-riding problem when the sampling size is small. When $s = 2$, forward-looking players who are aware of the effect of a single upward deviation predicted by sBRD might free ride from the very beginning if they believe there exist other FL players. As a result, they would play mBR and would not be identified as FL players. When the sampling size is large, the incentive to free ride is weakened as more upward deviations are required for successful efficiency-improving transitions.

3.3.4 Responsiveness to Signals

In this subsection, we explore a potential channel through which the sampling size affects the frequency of downward deviations. Subjects under the treatment $s = 2$ might consider the signals from only two random draws to be non-representative of the true population state, and thus they do not adjust their actions immediately even when a signal indicates action 2 (or 3) as the mBR. Therefore, a slow response to changes in the signals due to perception of inaccurate information could result in downward deviations. Here we test



Notes: Predicted probability is calculated based on the regression results reported in specification (1) of Table C.5. $Signal2$ and $Signal3$ are the fractions of action 2 and 3 in a signal respectively.

Figure. 3.7: Predicted probability of choosing each action given a signal.

whether subjects perceive signals with $s = 2$ to be less accurate than $s = 7$ so that they are less responsive to the changes in the signals (see Figure C.6 of Appendix C.2 for the subjects' action choices given each signal across treatments).

We run mixed logit regression, where the dependent variable is $Action_{it}$, the chosen action of a subject i in period t. The independent variables include the treatment of the sampling size $s7_i$ and the signals $Signal2_{it}$ and $Signal3_{it}$, which are the fraction of action 2 (3) observed in the signal by subject i in period t, respectively. The interaction terms of the sampling size and the signals $s7_i \times Signal2_{it}$ and $s7_i \times Signal3_{it}$ are included to test the effects of the the sampling size on the subjects' perceptions of signals' accuracy. The variables of personal characterises are added as controls. The standard errors are clustered at the session level.

Figure 3.7 illustrates the predicted probabilities of choosing each action given the fractions of actions 2 and 3, i.e., $Signal2$ and $Signal3$, respectively by treatment based on the results of regression using above specification (see specification (1) in Table C.5). It is clear that the probability of choosing an action increases with the fraction of the action in the signals. However, subjects' responsiveness to signals indeed differs across treatments - an increase in the fraction of an action in signals has a larger positive effect on the probability of the action to be chosen under $s = 7$ than $s = 2$ for all the three actions. (Also see the average marginal effects of the independent variables, reported in Table C.6 in Appendix C.2.)

Result 3.5. *Subjects under the smaller sampling size treatment $s = 2$ are less responsive to the changes in signals than those under the larger sampling size treatment $s = 7$.*

In Appendix C.1, we also discuss whether the inertia (i.e., the tendency to play the same action over time regardless of changes in signals) is an alternative channel that results in downward deviations, of which we find no supporting evidence.

3.3.5 Payoff Loss Dependent Deviations

In this subsection, we investigate how a deviation play is affected by the associated payoff loss, i.e., the expected payoff difference between the mBR action and the non-mBR action. Blume (1993); Myatt and Wallace (2003) propose a cost-dependent noise model where deviations involving a higher payoff loss are less likely to happen. However, the existing experimental results are mixed. Battalio et al. (2001); Mäs and Nax (2016); Hwang et al. (2018); Bilancini et al. (2020) find that deviations from mBR are cost-dependent, while Lim and Neary (2016) find that deviations are not cost-dependent. Our experimental results show that deviations are cost-dependent, but there is a significant difference in cost-dependence between deviations of different directions.

Figure 3.8 plots the the rate of deviation by direction and its associated payoff loss given each signal.[24] Note that both downward and upward deviation rates tend to decrease

[24] If there exist two possible actions for deviations, the one with a higher payoff is used, as the action associated with the lower payoff was chosen much less often - it consists of 1.71% of the data set and is not included for analysis. In Figure C.5 of Appendix C.2, we show that the results with the whole data set are qualitatively the same.

(a) Downward Deviation (b) Upward Deviation

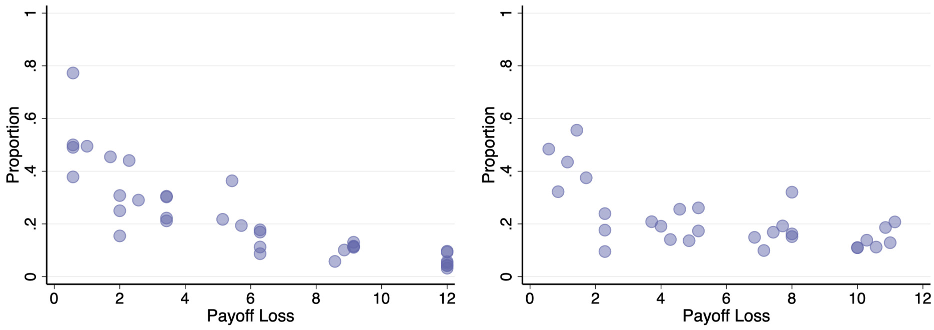


Figure. 3.8: Scatter plots of deviation rate and associated payoff loss given signals by direction.

with the payoff loss. Further, Spearman's rank order tests suggest a significant negative relationship between the deviation play and the payoff loss: Spearman's $\rho_{down} = -0.9072$, $p < 0.01$ for the downward deviation and $\rho_{up} = -0.5635$, $p < 0.01$ for the upward deviations.

More importantly, the test results suggest that the upward deviations are less correlated to the payoff loss than the downward deviations ($\Delta\rho = 0.3437$, $p = 0.037$). As we have shown, FL players are critical for transitions between strict Nash equilibria. These FL players are willing to bear some short-run losses to play upward deviations in order to obtain long-run gains. That is, upward deviations tend to be affected by payoff losses less than downward deviations.[25]

Result 3.6. *Deviation plays have a significant negative association with the expected payoff loss. Upward deviations are less sensitive to the payoff loss than downward deviations.*

[25]Restricting data to FL players gives Spearman's $\rho = -0.3697$, $p = 0.048$ for upward deviations. Though ρ is smaller than when all the players are included, the difference is not significant ($\Delta = -0.1938$, $p = 0.463$).

Table 3.8: Transition duration by session.

From e_1	Transition to e_2	Transition to e_3
Session 1, $s = 2$	-	36
Session 2, $s = 2$	-	20
Session 3, $s = 2$	60	-
Session 4, $s = 2$	9	-
Session 5, $s = 7$	-	6
Session 6, $s = 7$	-	74
Session 7, $s = 7$	-	71
Session 8, $s = 7$	-	65

3.3.6 Payoffs

In this subsection, we look at the realized payoffs in games. Transitions to the most efficient equilibrium e_3 are expected to increase payoffs. In the context of a coordination game, the payoff is also affected by the transition duration, i.e., the longer is the duration of the population state being unsettled, the lower the payoff will be. We define the transition duration from the initial state e_1 to the reached final equilibrium as the number of periods a population takes to reach and stay inside the basin of attraction of the equilibrium till the end of the game. For example, if a population state x stays in the basin of attraction of the final equilibrium reached since period $T + 1$ until the end of the game, that is, the population state x is not in the basin of attraction in period T, the transition duration is T.

Table 3.8 reports the transition durations from the initial state e_1 to e_2 or e_3 by session. The transition duration under $s = 7$ is on average longer than that under $s = 2$, regardless of the final equilibrium reached. Under $s = 7$, the four sessions transited to e_3 with the transition durations of 6, 74, 71, 65 periods, respectively. Under $s = 2$, sessions 1 and 2 transited to e_3 using 36 and 20 periods, respectively, and sessions 3 and 4 transited to e_2 using 60 and 9 periods, respectively. Thus, even though more sessions reach the most efficient equilibrium e_3 under $s = 7$, whether subjects can earn more under $s = 7$ than $s = 2$ is indeterminate because the transition duration under $s = 7$ is longer.

The average of accumulated individual total payoffs is shown in Figure 3.9 by treatment and the equilibrium reached. Let $\pi(s, e)$ denote the average accumulated individual payoff from groups that transited to equilibrium e under treatment s. We find that

$$\pi(2, e_3) > \pi(7, e_3) > \pi(2, e_2).$$

That is, subjects earned the highest payoff in groups that transited to e_3 under $s = 2$ ($\pi(2, e_3) = 2006.96$ tokens), which is significantly greater than those under $s = 7$ also transiting to e_3 ($\pi(2, e_3) = 1770.78$ tokens) by one-sided Mann-Whitney test ($\pi(2, e_3) > \pi(7, e_3)$, $p < 0.001$), while subjects in groups that transited to e_2 under $s = 2$ earned the lowest payoff ($\pi(2, e_2) = 1638.35$ tokens $< \pi(3, e_3)$ by one-sided Mann-Whitney test, $p = 0.072$).

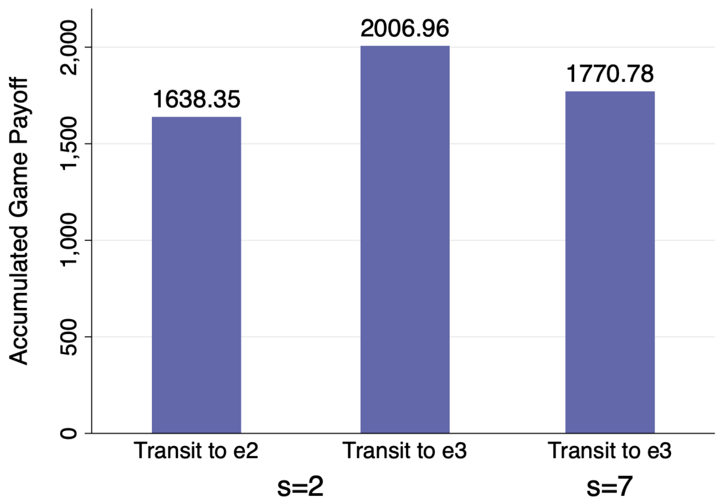


Figure. 3.9: Mean of realized individual payoffs (accumulated over 80 periods) by treatment and the final equilibrium reached.

The difference in the accumulated payoffs can be explained by transition durations and the final equilibrium reached. For example, the groups transited to the most

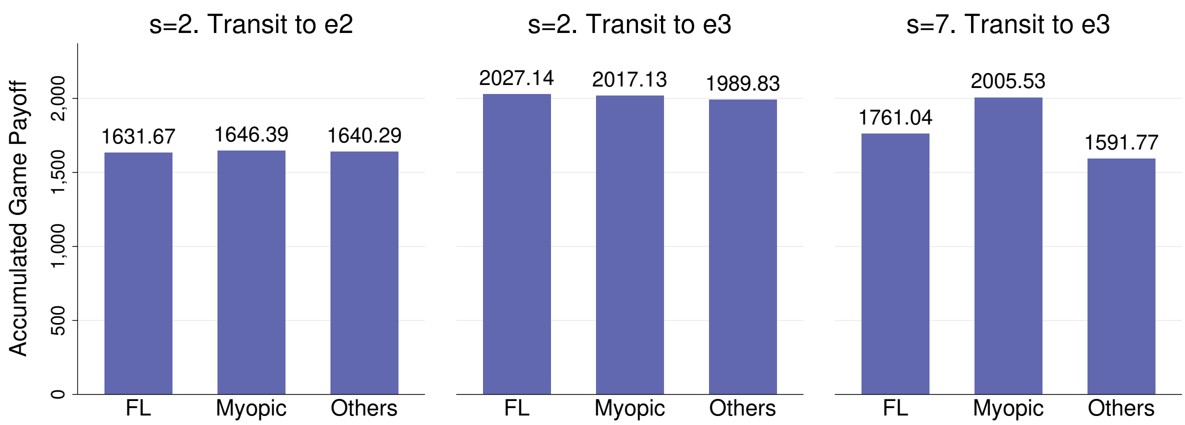


Figure. 3.10: Mean payoffs by subjects types.

efficient equilibrium e_3 under $s = 2$ spent less time on transitions than those under $s = 7$, suggesting that $\pi(2, e_3) > \pi(7, e_3)$. For the groups that transited to different equilibria, those transiting to e_3 have a higher payoff than those transiting to e_2, e.g., $\pi(2, e_3) > \pi(2, e_2)$, and $\pi(7, e_3) > \pi(2, e_2)$. The ranking in accumulated payoffs is robust to any longer time of the game play. If the game is played for another $n > 0$ periods, the accumulated payoffs will satisfy

$$\pi(2, e_3) + 30n > \pi(7, e_3) + 30n > \pi(2, e_2) + 20n,$$

where 30 and 20 are equilibrium payoffs of e_3 and e_2 respectively (see section 3.2.1).

Result 3.7. *The average accumulated individual payoffs in the game has the following ranking:*

$$\pi(2, e_3) > \pi(7, e_3) > \pi(2, e_2).$$

Result 3.7 has an interesting implication. On the one hand, more accurate information ensures transitions to the best equilibrium but involves a longer transition duration and a higher transition payoff loss. On the other hand, less accurate information involves a shorter transition duration but does not guarantee the final efficiency.

Figure 3.10 shows average payoffs by subjects' type. Under $s = 2$, all of the three types received similar payoffs regardless of the final equilibrium reached (Mann-Whitney test $p > 0.05$ for all of the pairwise comparisons). In contrast, under $s = 7$, the myopic subjects received the highest payoffs (2005.53), followed by the FL subjects (1761.04) and the rest subjects (1589.49) (Mann-Whitney test $p < 0.05$ for all of the pairwise comparisons).[26]

Result 3.8. *Under the treatment $s = 2$, the three types of subjects received similar payoffs regardless of the equilibria finally reached. Under the treatment $s = 7$, myopic subjects received significantly higher payoffs than FL subjects, followed by the others.*

This result has important policy implications on the tradeoff between welfare equality and efficiency when designing an information structure or institutions. For example, a small sampling size such as $s = 2$ can induce equal payoffs among subjects but it does not guarantee a transition to the more efficient equilibrium.

3.4 Conclusion

This paper provides an experimental investigation of the evolutionary game model (Oyama et al. (2015)) which predicts transitions among strict Nash equilibria under inexact information of opponents' behaviors. In our lab experiment, a population of 14 subjects played a three-action coordination game repeatedly for 80 periods with two informational treatments: the sampling size $s = 2$ for the less accurate information and $s = 7$ for the more accurate information. We observe that populations in all of the sessions transited away from the least efficient equilibrium under the two sampling sizes,

[26] Recall that subjects types are defined by patterns of deviation and mBR play (see section 3.3.3). Because deviations tend to involve payoff loss, one might expect payoffs to differ by type under both treatments. The equal payoff across types under $s = 2$ could be explained by the small sampling size with which the signals are less representative to the true population state, thus mBR is less likely to be empirically optimal (and a deviation is more likely to be profitable). See Figure C.8 for payoffs over time by session.

but more sessions transited to the most efficient equilibrium under $s = 7$ than $s = 2$, which is in contrast with the theory.

We find that the larger sampling size $s = 7$ induces significantly more subjects to strategically play upward deviations that can lead the population to a more efficient equilibrium, although there is no significant treatment effect on the probability of such upward deviation play. On the other hand, the failure of transitions to the efficient equilibrium under $s = 2$ arises from the fact that subjects are less elastic to changes in their private signals due to the less accurate signals, leading to significantly more downward deviations, which either delay or block the transitions.

Our results on the equilibrium transitions and the underlying individual behaviors suggest that human subjects tend to disregard inaccurate signals despite the fact that the signals are unbiased. More importantly, human subjects are more likely to exhibit farsighted strategic behaviors when information is relatively more accurate. One insight from our results is that researchers and policy-makers might take these two types of behaviors into consideration when designing any information-related mechanisms or institutions.

Finally, we find that although accurate information can induce transitions to the efficient equilibrium, it also involves a higher welfare loss during the transitions. In our experiment, subjects under $s = 7$ did not earn significantly more payoffs than $s = 2$ due to longer transition durations. Moreover, under $s = 7$, myopic subjects received higher payoffs than forward-looking subjects, followed by others; while under $s = 2$ three types earned equal amounts. Therefore, a small sampling size might be desirable when equality is preferred over efficiency.

Bibliography

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman, "Beliefs and Endogenous Cognitive Levels: an Experimental Study," *Games and Economic Behavior*, 2012, *75* (2), 449–463.

Aguinis, Herman, Isabel Villamor, and Ravi S. Ramani, "MTurk Research: Review and Recommendations," *Journal of Management*, 2021, *47* (4), 823–837.

Alan, Sule and Seda Ertac, "Fostering Patience in the Classroom: Results from Randomized Educational Intervention," *Journal of Political Economy*, 2018, *126* (5), 1865–1911.

_ , Teodora Boneva, and Seda Ertac, "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit," *The Quarterly Journal of Economics*, 2019, *134* (3), 1121–1162.

Alaoui, Larbi and Antonio Penta, "Endogenous Depth of Reasoning," *The Review of Economic Studies*, 2016, *83* (4 (297)), 1297–1333.

_ , Katharina A. Janezic, and Antonio Penta, "Reasoning about Others' Reasoning," *Journal of Economic Theory*, 2020, *189*, 105091.

Allgood, Sam, William B. Walstad, and John J. Siegfried, "Research on Teaching Economics to Undergraduates," *Journal of Economic Literature*, June 2015, *53* (2), 285–325.

Amir, Ofra, David G Rand, and Ya'akov Kobi Gal, "Economic Games on the Internet: The Effect of $1 Stakes," *PloS one*, 2012, *7* (2), e31461.

Anctil, Regina M, John Dickhaut, Chandra Kanodia, and Brian Shapiro, "Information Transparency and Coordination Failure: Theory and Experiment," *Journal of Accounting Research*, 2004, *42* (2), 159–195.

Andrade, Heidi L., "A Critical Review of Research on Student Self-Assessment," *Frontiers in Education*, 2019, *4*, 87.

Andreoni, James, Nikos Nikiforakis, and Simon Siegenthaler, "Predicting Social Tipping and Norm Change in Controlled Experiments," *Proceedings of the National Academy of Sciences*, 2021, *118* (16).

Arad, Ayala and Ariel Rubinstein, "The 11-20 Money Request Game: A Level-k Reasoning Study," *American Economic Review*, 2012, *102* (7), 3561–73.

Arechar, Antonio A, Simon Gächter, and Lucas Molleman, "Conducting Interactive Experiments Online," *Experimental economics*, 2018, *21*, 99–131.

Battalio, Raymond, Larry Samuelson, and John Van Huyck, "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games," *Econometrica*, 2001, *69* (3), 749–764.

Berinsky, Adam J, Gregory A Huber, and Gabriel S Lenz, "Evaluating Online Labor Markets for Experimental Research: Amazon. com's Mechanical Turk," *Political analysis*, 2012, *20* (3), 351–368.

Bilancini, Ennio, Leonardo Boncinelli, and Heinrich H Nax, "What Noise Matters? Experimental Evidence for Stochastic Deviations in Social Norms," *Journal of Behavioral and Experimental Economics*, 2020, *90*, 101626.

Blume, Lawrence E, “The Statistical Mechanics of Strategic Interaction,” *Games and Economic Behavior*, 1993, *5* (3), 387–424.

Bosch-Rosa, Ciril and Thomas Meissner, “The One Player Guessing Game: A Diagnosis on the Relationship between Equilibrium Play, Stated Beliefs, and Best Responses,” *Experimental Economics*, 2020, *23* (4), 1129–1147.

Brañas-Garza, Pablo, Teresa García-Muñoz, and Roberto Hernán González, “Cognitive Effort in the Beauty Contest Game,” *Journal of Economic Behavior & Organization*, 2012, *83* (2), 254–260.

Burnham, Terence C., David Cesarini, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace, “Higher Cognitive Ability is Associated with Lower Entries in a p-Beauty Contest,” *Journal of Economic Behavior & Organization*, 2009, *72* (1), 171–175.

Cabrales, Antonio, Rosemarie Nagel, and Roc Armenter, “Equilibrium Selection Through Incomplete Information in Coordination Games: An Experimental Study,” *Experimental Economics*, 2007, *10* (3), 221–234.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong, “Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games,” *Journal of Economic Theory*, 2002, *104* (1), 137–188.

_ , _ , and _ , “A Cognitive Hierarchy Model of Games,” *The Quarterly Journal of Economics*, 08 2004, *119* (3), 861–898.

Cappelen, Alexander, John List, Anya Samek, and Bertil Tungodden, “The Effect of Early-Childhood Education on Social Preferences,” *Journal of Political Economy*, 2020, *128* (7), 2739–2758.

Capra, C Mónica, "Understanding Decision Processes in Guessing Games: A Protocol Analysis Approach," *Journal of the Economic Science Association*, 2019, *5* (1), 123–135.

Chen, Daniel L., Martin Schonger, and Chris Wickens, "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Cornand, Camille, "Speculative Attacks and Informational Structure: An Experimental Study," *Review of International Economics*, 2006, *14* (5), 797–817.

Costa-Gomes, Miguel A. and Vincent P. Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 2006, *96* (5), 1737–1768.

Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri, "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, March 2013, *51* (1), 5–62.

Ditzen, Jan, Yiannis Karavias, and Joakim Westerlund, "Testing and Estimating Structural Breaks in Time Series and Panel Data in Stata," 2021. Working Paper.

Dominitz, Jeff, "Earnings Expectations, Revisions, and Realizations," *Review of Economics and Statistics*, 1998, *80* (3), 374–388.

Duffy, John and Jonathan Lafky, "Social Conformity Under Evolving Private Preferences," *Games and Economic Behavior*, 2021, *128*, 104–124.

_ , Ed Hopkins, and Tatiana Kornienko, "Facing the Grim Truth: Repeated Prisoner's Dilemma Against Robot Opponents," 2022. Working Paper.

Dunning, David, Chip Heath, and Jerry M. Suls, "Flawed Self-Assessment: Implications for Health, Education, and the Workplace," *Psychological Science in the Public Interest*, 2004, *5* (3), 69–106.

Eckel, Catherine C. and Philip J. Grossman, "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk," *Evolution and Human Behavior*, 2002, *23* (4), 281–295.

Ellison, Glenn, "Learning from personal experience: One rational guy and the justification of myopia," *Games and Economic Behavior*, 1997, *19* (2), 180–210.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *The Quarterly Journal of Economics*, 05 2018, *133* (4), 1645–1692.

_ , _ , _ , David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," 2016. IZA Discussion Paper No. 9674.

Fan, Chinn-Ping, "Teaching Children Cooperation — An Application of Experimental Game Theory," *Journal of Economic Behavior & Organization*, 2000, *41* (3), 191–209.

Ferraro, Paul J., "Know Thyself: Competence and Self-awareness," *Atlantic Economic Journal*, 2010, *38* (2), 183–196.

Fischer, Mira and Dirk Sliwka, "Confidence in Knowledge or Confidence in the Ability to Learn: An Experiment on the Causal Effects of Beliefs on Motivation," *Games and Economic Behavior*, 2018, *111*, 122–142.

Fischhoff, Baruch, Aandrew M Parker, Wändi Bruine de Bruin, Julie Downs, Claire Palmgren, Robyn Dawes, and Charles F. Manski, "Teen Expectations for Significant Life Events," *The Public Opinion Quarterly*, 2000, *64* (2), 189–205.

Foster, Dean and Peyton Young, "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 1990, *38* (2), 219–232.

Franchetti, Francisco and William H. Sandholm, "An Introduction to Dynamo: Diagrams for Evolutionary Game Dynamics," *Biological Theory*, 2013, *8* (2), 167–178.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, December 2005, *19* (4), 25–42.

Friedenberg, Amanda, Willemien Kets, and Terri Kneeland, "Is Bounded Reasoning about Rationality Driven by Limited Ability?," 2021. Working Paper.

Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber, "On the Persistence of Strategic Sophistication," *Journal of Economic Theory*, 2015, *159*, 369–400.

Gilboa, Itzhak and Akihiko Matsui, "Social Stability and Equilibrium," *Econometrica*, 1991, *59* (3), 859–867.

Gill, David and Victoria Prowse, "Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis," *Journal of Political Economy*, 2016, *124* (6), 1619–1676.

Goodman, Joseph K., Cynthia E. Cryder, and Amar Cheema, "Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples," *Journal of Behavioral Decision Making*, 2013, *26* (3), 213–224.

Grimes, Paul W., "The Overconfident Principles of Economics Student: An Examination of a Metacognitive Skill," *The Journal of Economic Education*, 2002, *33* (1), 15–30.

Grossman, Zachary and David Owens, "An Unlucky Feeling: Overconfidence and Noisy Feedback," *Journal of Economic Behavior & Organization*, 2012, *84* (2), 510–524.

Guest, Jon and Robert Riegler, “Learning by Doing: Do Economics Students Self-Evaluation Skills Improve?,” *International Review of Economics Education*, 2017, *24*, 50–64.

Halevy, Yoram, Johannes Hoelzemann, and Terri Kneeland, “Magic Mirror on the Wall, Who Is the Smartest One of All?,” 2021. Working Paper.

Hauser, David, Gabriele Paolacci, and Jesse Chandler, “Common Concerns with MTurk as a Participant Pool: Evidence and Solutions,” in “Handbook of research methods in consumer psychology,” Routledge/Taylor & Francis Group, 2019, pp. 319–337.

Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels, “The Theory of Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information,” *Econometrica*, 2004, *72* (5), 1583–1599.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt, “Iterated Dominance and Iterated Best Response in Experimental “p-Beauty Contests",” *American Economic Review*, 1998, *88* (4), 947–969.

Horton, John J, David G Rand, and Richard J Zeckhauser, “The Online Laboratory: Conducting Experiments in a Real Labor Market,” *Experimental Economics*, 2011, *14*, 399–425.

Hossain, Belayet and Panagiotis Tsigaris, “Are Grade Expectations Rational? A Classroom Experiment,” *Education Economics*, 2015, *23* (2), 199–212.

Huyck, John Van, Ajalavat Viriyavipart, and Alexander L. Brown, “When Less Information is Good Enough: Experiments with Global Stag Hunt Games,” *Experimental Economics*, 2018, *21* (3), 527–548.

Hwang, Sung-Ha, Wooyoung Lim, Philip Neary, and Jonathan Newton, "Conventional Contracts, Intentional Behavior and Logit Choice: Equality without Symmetry," *Games and Economic Behavior*, 2018, *110*, 273–294.

James, Harvey S. and Jeffrey P. Cohen, "Does Ethics Training Neutralize the Incentives of the Prisoner's Dilemma? Evidence from a Classroom Experiment," *Journal of Business Ethics*, 2004, *50* (1), 53–61.

Jin, Ye, "Does Level-k Behavior Imply Level-k thinking?," *Experimental Economics*, 2021, *24* (1), 330–353.

_ , "Reinvestigating Rk Behavior in Ring Games," *Journal of Behavioral and Experimental Economics*, 2022, *98*, 101878.

Johnson, Eric J., Colin Camerer, Sankar Sen, and Talia Rymon, "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining," *Journal of Economic Theory*, 2002, *104* (1), 16–47.

Kandori, Michihiro, George J Mailath, and Rafael Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 1993, pp. 29–56.

Kneeland, Terri, "Identifying Higher-Order Rationality," *Econometrica*, 2015, *83* (5), 2065–2079.

Kruger, Justin and David Dunning, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, 1999, *77* (6), 1121.

Lim, Wooyoung and Philip R. Neary, "An Experimental Investigation of Stochastic Adjustment Dynamics," *Games and Economic Behavior*, 2016, *100*, 208–219.

Lyu, Jianxun, “Dynamic Cognitive Hierarchy Model: A Theory of Equilibrium Switching,” 2022. Working Paper.

Martins, Julia, Richard Fortmüller, and Owen Powell, “Can Teaching Social Dilemmas Make People More Prosocial? An Experiment,” *Journal of Education for Business*, 2017, *92* (1), 16–22.

Mäs, Michael and Heinrich H Nax, “A Behavioral Study of “Noise” in Coordination Games,” *Journal of Economic Theory*, 2016, *162*, 195–208.

Moore, Don A. and Derek Schatz, “The Three Faces of Overconfidence,” *Social and Personality Psychology Compass*, 2017, *11* (8), e12331.

_ **and Paul J. Healy**, “The Trouble with Overconfidence,” *Psychological Review*, 2008, *115* (2), 502.

Myatt, David P and Chris Wallace, “A Multinomial Probit Model of Stochastic Evolution,” *Journal of Economic Theory*, 2003, *113* (2), 286–301.

Nagel, Rosemarie, “Unraveling in Guessing Games: An Experimental Study,” *American Economic Review*, 1995, *85* (5), 1313–1326.

Naidu, Suresh, Sung-Ha Hwang, and Samuel Bowles, “Evolutionary Bargaining with Intentional Idiosyncratic Play,” *Economics Letters*, 2010, *109* (1), 31–33.

Nowell, Clifford and Richard M. Alston, “I Thought I Got an A! Overconfidence Across the Economics Curriculum,” *The Journal of Economic Education*, 2007, *38* (2), 131–142.

Oyama, Daisuke, William H. Sandholm, and Olivier Tercieux, “Sampling Best Response Dynamics and Deterministic Equilibrium Selection,” *Theoretical Economics*, 2015, *10* (1), 243–281.

Palacios-Huerta, Ignacio and Oscar Volij, "Field Centipedes," *American Economic Review*, September 2009, *99* (4), 1619–35.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis, "Running Experiments on Amazon Mechanical Turk," *Judgment and Decision making*, 2010, *5* (5), 411–419.

Puente, Manuel, "Disentangling Experimental Evidence on Higher Order Rationality," 2020. Working Paper.

Reynaud, Arnaud and Stéphane Couture, "Stability of Risk Preference Measures: Results from a Field Experiment on French Farmers," *Theory and Decision*, 2012, *73* (2), 203–221.

Sandholm, William H, "Almost Global Convergence to p-dominant Equilibrium," *International Journal of Game Theory*, 2001, *30* (1), 107–116.

Sawa, Ryoji and Jiabin Wu, "Statistical Inference in Evolutionary Dynamics," *Games and Economic Behavior*, 2023, *137*, 294–316.

Schlösser, Thomas, David Dunning, Kerri L. Johnson, and Justin Kruger, "How Unaware are the Unskilled? Empirical Tests of the "Signal Extraction" Counter-explanation for the Dunning–Kruger Effect in Self-Evaluation of Performance," *Journal of Economic Psychology*, 2013, *39*, 85–100.

Schnusenberg, Oliver and Andrés Gallo, "On Cognitive Ability And Learning In A Beauty Contest," *Journal for Economic Educators*, 2011, *11* (1), 13–24.

Siegfried, John J. and William B. Walstad, "Research on Teaching College Economics," in William B. Walstad and Phillip Saunders, eds., *Teaching Undergraduate Economics: A Handbook for Instructor*, New York: Irwin/McGraw-Hill, 1998, pp. 141–66.

Simon, Herbert A., "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 1955, *69* (1), 99–118.

— , "From Substantive to Procedural Rationality," in "25 Years of Economic Theory," Springer, 1976, pp. 65–86.

— , *Models of Bounded Rationality*, Cambridge Mass: MIT Press, 1982.

Slonim, Robert L., "Competing Against Experienced and Inexperienced Players," *Experimental Economics*, 2005, *8* (1), 55–75.

Smerdon, David, Theo Offerman, and Uri Gneezy, "Everybody's Doing It: On the Persistence of Bad Social Norms," *Experimental Economics*, 2019, pp. 1–29.

Snowberg, Erik and Leeat Yariv, "Testing the Waters: Behavior across Participant Pools," *American Economic Review*, 2021, *111* (2), 687–719.

Spooren, Pieter, Bert Brockx, and Dimitri Mortelmans, "On the Validity of Student Evaluation of Teaching: The State of the Art," *Review of Educational Research*, 2013, *83* (4), 598–642.

Sprenger, Charles and Wanqun Zhao, "Predictive Accuracy for Measures of Higher Order Rationality - An Experiment Using Ring Games," 2021. Working Paper.

Stahl, Dale O. and Paul W. Wilson, "Experimental Evidence on Players' Models of Other Players," *Journal of economic behavior & organization*, 1994, *25* (3), 309–327.

— **and** — , "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 1995, *10* (1), 218–254.

Thoma, Carmen, "Under- versus Overconfidence: An Experiment on How Others Perceive a Biased Self-Assessment," *Experimental Economics*, 2016, *19* (1), 218–239.

Toplak, Maggie E., Richard F. West, and Keith E. Stanovich, "The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-Biases Tasks," *Memory & Cognition*, 2011, *39* (7), 1275.

Verbrugge, Rineke, Ben Meijering, Stefan Wierda, Hedderik van Rijn, and Niels Taatgen, "Stepwise Training Supports Strategic Second-order Theory of Mind in Turn-taking Games," *Judgment and Decision Making*, 2018, *13* (1), 79–98.

Wason, Peter C., "Reasoning About a Rule," *Quarterly Journal of Experimental Psychology*, 1968, *20* (3), 273–281.

Weber, Roberto A., "'Learning' with no Feedback in a Competitive Guessing Game," *Games and Economic Behavior*, 2003, *44* (1), 134–144.

Young, H. Peyton, "The Evolution of Conventions," *Econometrica*, 1993, pp. 57–84.

Zhou, Zhengyuan, Susan Athey, and Stefan Wager, "Offline Multi-Action Policy Learning: Generalization and Optimization," *Operations Research*, 2022.

Appendix A for Chapter 1

A.1 Time Spent on the Tutorials

In this section, I examine a possibility of a negative impact of the color tutorial on performance in Part 3: subjects in the control group had to switch attention between game irrelevant and relevant tasks, possibly inducing higher working memory and cognitive load when they played games in Part 3. If this is the case, there could be a confound of the significant treatment effect on Ak_3 that is found in the main paper.

I use time spent on a task (or response time) as a measure of cognitive load, and assume that the higher cognitive load induces a longer response time. On average the control group spent significantly more time (minutes) on the tutorial than the treatment group (6.794 vs. 6.105, Mann–Whitney test $p < 0.01$). This group difference in time spent on the tutorials (tutorial time hereafter) can also be seen from column (3) of Table A.1 which reports tobit regression of the tutorial time on treatment, controlling for pre-tutorial ability Ak_1, CRT and demographics. However, there is a possibility that the time difference is driven by economics students in the treatment group who spent significantly less time on the relevant tutorial (column (2)), possibly due to their prior exposure to similar economic concepts. Indeed, column (4) with the interaction of treatment and economics shows that economics students in the treatment groups spent less time on the tutorial at 10% significant level. In contrast, the treatment on its own is insignificant, suggesting that the cognitive loads of the two tutorials are similar to non-economics students.

More importantly, the tutorial time does not significantly decrease post-tutorial reasoning ability either in the control group, the treatment group or the two groups pooled together, when pre-tutorial ability and personal characteristics are controlled for (columns (1), (3) and (5) of Table A.2). Furthermore, adding the interaction of economics and tutorial time does not alter the results. In particular, neither tutorial time on its own nor the interaction term significantly predicts post-tutorial performance (columns (2), (4), (6)).

Result A.1. *There is no evidence that the cognitive load of the irrelevant tutorial worsens subjects' post-tutorial ability, and thus it would not confound with the treatment effect on the reasoning ability.*

Table A.1: Time spent on the tutorials.

Tutorial Time	(1) Control	(2) Treatment	(3) Total	(4) Total
Treatment			-0.544**	-0.199
			(0.237)	(0.311)
Treatment×Economics				-0.876*
				(0.444)
Ak_1	-0.141	-0.130	-0.122	-0.135
	(0.145)	(0.084)	(0.079)	(0.084)
CRT	0.053	-0.183*	-0.081	-0.074
	(0.177)	(0.098)	(0.099)	(0.103)
Female	0.328	0.311	0.325	0.310
	(0.370)	(0.338)	(0.256)	(0.250)
Economics	0.278	-0.575***	-0.178	0.274
	(0.475)	(0.160)	(0.273)	(0.414)
Age	0.001	0.004	0.004	0.002
	(0.024)	(0.014)	(0.014)	(0.014)
Study Level	0.474***	0.318*	0.391***	0.395***
	(0.170)	(0.186)	(0.127)	(0.130)
Constant	5.332***	6.032***	5.957***	5.827***
	(1.145)	(0.673)	(0.721)	(0.763)
N	76	79	155	155

Notes: Tutorial time = time spent on the tutorials. Tobit regressions with lower bound = 5 minutes and upper bound = 10 minutes. Session clustered standard error in the parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: The effect of time spent on the tutorial on post-tutorial ability.

Ak_3	Control		Treatment		Total	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment					0.892**	0.849*
					(0.439)	(0.450)
Tutorial Time	0.220	0.230	-0.402	-0.305	-0.152	-0.089
	(0.145)	(0.221)	(0.252)	(0.296)	(0.156)	(0.201)
Economics	0.127	0.293	0.983*	2.933*	0.713*	1.829
	(0.294)	(1.912)	(0.589)	(1.775)	(0.381)	(1.393)
Economics×Tutorial Time		-0.024		-0.329		-0.172
		(0.298)		(0.295)		(0.200)
Ak_1	1.201***	1.199***	0.564***	0.557***	0.648***	0.643***
	(0.293)	(0.300)	(0.210)	(0.213)	(0.138)	(0.143)
CRT	0.066	0.068	0.706***	0.731***	0.457***	0.465***
	(0.162)	(0.181)	(0.159)	(0.186)	(0.108)	(0.112)
Female	-0.646	-0.643	0.350	0.362	-0.096	-0.088
	(0.556)	(0.541)	(0.750)	(0.735)	(0.419)	(0.413)
Age	-0.032	-0.032	0.018	0.017	0.000	0.001
	(0.023)	(0.022)	(0.033)	(0.032)	(0.024)	(0.024)
Study Level	-0.211	-0.214	0.366	0.307	0.125	0.104
	(0.407)	(0.420)	(0.467)	(0.479)	(0.280)	(0.278)
N	76	76	79	79	155	155

Notes: Ordered logit. Session clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.2 Additional Tables and Figures

Table A.3: Summary Statistics. Mean/Share (SD).

	Control	Treatment	Total
Female	0.618	0.506	0.561
	(0.489)	(0.503)	(0.498)
Age	24.05	23.56	23.80
	(8.142)	(8.080)	(8.088)
Study Level	2.468	2.263*	2.363
	(0.754)	(0.689)	(0.726)
Economics	0.364	0.438	0.401
	(0.484)	(0.499)	(0.492)
CRT	2.494	2.737	2.618
	(1.242)	(1.300)	(1.274)

Notes: Study Level = 1 if not a student; 2 if a undergraduate; 3 if a taught postgraduate; 4 if a research postgraduate. See Toplak et al. (2011) for the four-item CRT questions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ according to Mann–Whitney test of difference in each variable between control and treatment group.

Table A.4: Subjects' types before and after the tutorials (subjects are assigned to the higher type if subjects can be classified into two types with an error). Counts.

		Control: Irrelevant Tutorial						Treatment: Relevant Tutorial					
		Part 3: After Tutorial						Part 3: After Tutorial					
		A0	A1	A2	A3	A4	Total	A0	A1	A2	A3	A4	Total
Part 1: Before Tutorial	A0	16	9	6	1	2	34	13	2	1	1	18	35
	A1	2	14	3	2	0	21	5	9	1	2	5	22
	A2	4	3	2	1	0	10	4	0	0	0	7	11
	A3	0	0	2	0	0	2	0	0	0	0	2	2
	A4	0	0	0	1	9	10	0	0	0	0	10	10
	Total	22	26	13	5	**11**	77	22	11	2	3	**42**	80

Notes: The shades of the cells represents how ability changes from Part 1 to Part 3: light - lower ability; medium - higher ability; dark - same ability.

Table A.5: Subjects' types before and after the tutorials (two errors). Counts.

		Control: Irrelevant Tutorial Part 3: After Tutorial						Treatment: Relevant Tutorial Part 3: After Tutorial					
		A0	A1	A2	A3	A4	Total	A0	A1	A2	A3	A4	Total
Part 1: Before Tutorial	A0	5	6	1	0	1	13	6	3	3	0	6	18
	A1	3	25	7	1	2	38	5	11	1	5	17	39
	A2	2	7	4	1	0	14	0	2	1	0	8	11
	A3	0	0	1	1	0	2	0	0	0	0	4	4
	A4	0	0	0	2	8	10	0	0	0	0	8	8
	Total	10	38	13	5	**11**	77	11	16	5	5	**43**	80

Notes: The shades of the cells represents how ability changes from Part 1 to Part 3: light - lower ability; medium - higher ability; dark - same ability.

Table A.6: Alternative specifications for testing treatment effect on ability improvement.

$Ak_3 > Ak_1$	(1)	(2)	(3)	(4)
Treatment	0.985***	0.899**	0.822**	-1.157
	(0.356)	(0.378)	(0.418)	(0.827)
CRT		0.383***	0.362**	-0.064
		(0.146)	(0.144)	(0.199)
Treatment×CRT				0.786***
				(0.265)
Economics			0.395	0.327
			(0.485)	(0.500)
Female			-0.273	-0.276
			(0.404)	(0.411)
Study Level			0.089	0.078
			(0.259)	(0.289)
Age			-0.021	-0.023
			(0.027)	(0.028)
N	141	141	139	139

Notes: Logit. $Ak_3 > Ak_1$ is an indicator for subjects making ability improvement. Session clustered standard error in the parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $A4_1$ excluded.

Table A.7: Relationship between treatment, and estimated performance.

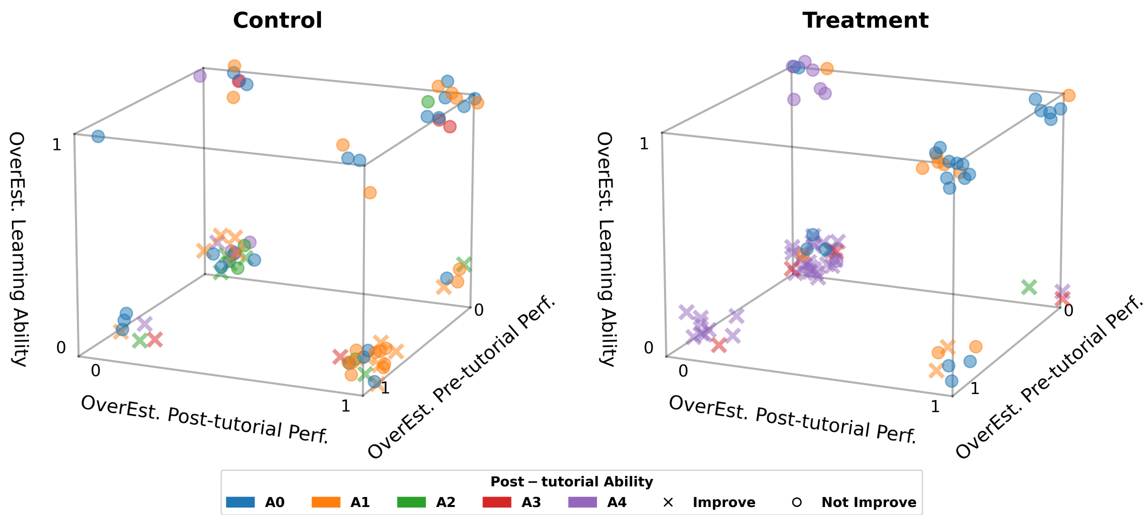
	$\hat{x}_1$			$\hat{x}_3$		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.366	-0.530*	-0.733	1.227***	1.286***	0.451
	(0.289)	(0.300)	(0.562)	(0.233)	(0.238)	(0.372)
$Ak_3 > Ak_1$	-0.335	-0.337	-0.570	0.753	0.637	-0.367
	(0.381)	(0.382)	(0.471)	(0.513)	(0.478)	(0.495)
Treatment×($Ak_3 > Ak_1$)			0.455			2.100**
			(0.797)			(0.943)
CRT		-0.043	-0.063		0.144	0.060
		(0.142)	(0.141)		(0.146)	(0.147)
Female		-0.138	-0.155		-0.216	-0.313
		(0.272)	(0.269)		(0.398)	(0.407)
Economics		0.516	0.474		1.105***	0.943***
		(0.356)	(0.363)		(0.265)	(0.326)
Study Level		-0.338	-0.330		-0.184	-0.179
		(0.336)	(0.347)		(0.233)	(0.252)
Age		0.031	0.030		0.021	0.015
		(0.020)	(0.020)		(0.018)	(0.017)
N	141	139	139	141	139	139

Notes: Test of treatment effect on estimations, and robustness check of treatment effect on perceived performance improvement shown in Table 1.8. Ordered logit. Session clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $A4_1$ excluded.

Table A.8: Treatment and distance between estimation and performance $|\hat{x} - x|$.

	$|\hat{x}_1 - x_1|$				$|\hat{x}_3 - x_3|$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.153	0.175	0.153	-0.097	-0.722**	-0.466	-0.545*	0.115
	(0.316)	(0.335)	(0.357)	(0.384)	(0.300)	(0.287)	(0.288)	(0.198)
Ak_1		0.057	0.082	0.402**		-0.192*	-0.119	-0.226
		(0.175)	(0.170)	(0.190)		(0.113)	(0.145)	(0.205)
Ak_3		-0.013	-0.084	-0.189		-0.446***	-0.393***	-0.082
		(0.067)	(0.084)	(0.206)		(0.112)	(0.128)	(0.175)
$A4_3$				1.476*				0.999
				(0.806)				(0.990)
Treatment$\times A4_3$				-0.992**				-3.085***
				(0.439)				(1.039)
CRT			0.104	0.118			-0.315*	-0.287
			(0.138)	(0.124)			(0.187)	(0.198)
Female			0.237	0.328			0.128	0.384
			(0.246)	(0.253)			(0.315)	(0.379)
Economics			0.365	0.434			0.252	0.334
			(0.373)	(0.407)			(0.351)	(0.474)
Age			0.033	0.039			0.030*	0.034
			(0.026)	(0.026)			(0.017)	(0.023)
Study Level			-0.343*	-0.481**			0.058	0.088
			(0.183)	(0.216)			(0.224)	(0.233)
N	157	157	155	139	157	157	155	139

Notes: Ordered Logit. Session clustered Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns (4) and (8): $A4_1$ excluded.



Notes: subjects who are $A4$ in Part 1 are place in the category 'A4 Not Improve' because there is no space for improvement for them.

Figure. A.1: Scatter plot of estimation of performance.

A.3 Instructions

Instructions for Part 1

Overview

Thanks again for your participation!

The experiment consists of 5 parts. You will be provided with instructions for Part 1 first. To ensure that you fully understand the instructions for Part 1, you will asked to take a comprehension quiz which will be shown on the computer screen after you finish reading the instructions. While you are answering the quiz, please feel free to refer to the paper instructions. After Part 1 is complete, you will be provided with separate instructions before each further part begins. If you have any questions about the instructions prior to the start of each part, please raise your hand so that an experimenter would come to help you in person.

Details

This part of the experiment has 8 rounds. In each round, you will be randomly assigned a role of a player against three other players in a 4-player game, where each player's earnings will depend on the combination of their own action and one other player's action. These earnings possibilities will be represented in payoff tables like in Figure 1, where the first table starting from the left side is the payoff table for Player 1 whose payoff depends on the actions taken by both Player 1 and Player 2. The second table represents the payoffs for Player 2 whose payoff depends on the actions taken by both Player 2 and Player 3. The third table is the payoff table for Player 3 whose payoff depends on the actions of both Player 3 and Player 4. The last table is the payoff table for Player 4 whose payoff depends on the actions by both Player 4 and Player 1.

Player 1's Payoffs

Player 1' actions \ Player 2's actions	d	e	f
a	10	4	16
b	20	8	0
c	4	18	12

Player 2's payoffs

Player 2' actions \ Player 3's actions	g	h	i
d	12	16	4
e	0	12	8
f	4	4	20

Player 3's payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	20	12	8
h	6	8	18
i	0	16	4

Player 4's payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	10	12	8
k	6	20	18
l	16	4	0

Figure. A.2: Example 1.

In each round, you will be asked to choose one of three available actions. Each other player in your game will also choose one of three actions. The actions that you could choose from are labelled in the column with letters in the payoff tables, while the actions available for your opponent are labelled in the row with letters.

For example, suppose you were assigned a role of Player 1 in the game represented in Figure 1. Then your payoff table is given by the leftmost table, and your available actions and your possible payoffs (which are jointly determined by your action and Player 2's action) are highlighted in red. You should choose among a, b, and c while Player 2, who is your opponent, will choose among d, e, and f.

Your earnings will depend jointly on your action and your opponent's actions. Your action (either a, b, or c) will determine the row of the leftmost table and your opponent Player 2's action (either d, e, or f) will determine the column of the table. The cell corresponding to this combination of actions will determine your earnings. (You may want to move your mouse over a table cell to see the associated action of each player as the corresponding row and column will be highlighted in blue; you may also want to highlight a table cell in orange by clicking it; to undo highlighting, click the cell again.) For example, if you choose action a, and Player 2 chooses action d, your payoff is 10 tokens; if instead Player 2 chooses *e*, you would earn 4 tokens, and if instead Player 2 chooses *f*, you would earn 16 tokens. And if you choose action *b* and Player 2 chooses *f*, your payoff is 0 tokens; if you choose *c*, Player 2 chooses *e*, your payoff is 18 tokens, and so on.

Player 2, Player 3, and Player 4's actions and earnings are listed in the other three payoff tables. Player 2 may choose either *d*, *e*, or *f*, while Player 3 may choose either *g*, *h*, or *i*, and Player 4 may choose either *j*, *k*, or *l*. Player 2's earnings depend jointly upon the action he chooses and the action Player 3 chooses. Player 3's earnings depend jointly upon the action he chooses and the action Player 4 chooses. Player 4's earnings jointly depend upon the action he chooses and the action Player 1 chooses. So if you, as Player 1, choose *c*, Player 2 chooses *f*, Player 3 chooses *g*, Player 4 chooses *j*, then your payoff is 12 tokens, Player 2's payoff is 4 tokens, Player 3's payoff is 20 tokens and Player 4's payoff is 8 tokens.

So if, instead, your role is Player 2, you should choose among *d*, *e*, and *f*, while your opponent Player 3 will choose among *g*, *h*, and *i*, and so on. If you, as Player 2, choose *d*, Player 3 chooses *i*, Player 4 chooses *l*, and Player 1 chooses *c*, then your payoff is 4 tokens, Player 3's payoff is 4 tokens, Player 4's payoff is 0 tokens and Player 1's payoff is 4 tokens.

The different earnings tables will appear in a random order in each round. As well, the earnings tables and your role will differ from game to game. So you should always look at the role assigned to you and order of the tables carefully at the beginning of each game. For example, the payoff tables may look like in Figure 2 instead of the example in Figure 1, and you might instead be assigned to the role of Player 3, whose payoff table is highlighted in red.

Importantly, in each round, you will be matched to play against 3 *computer players*. For example, if you are Player 1, the other 3 players are all computer players. Note that

Player 2's Payoffs

Player 2' s actions \ Player 3's actions	g	h	i
d	12	16	4
e	0	12	8
f	4	4	20

Player 4's payoffs

Player 4' s actions \ Player 1's actions	a	b	c
j	10	12	8
k	6	20	18
l	16	4	0

Player 1's payoffs

Player 1' s actions \ Player 2's actions	d	e	f
a	10	4	16
b	20	8	0
c	4	18	12

Player 3's payoffs

Player 3' s actions \ Player 4's actions	j	k	l
g	36	12	8
h	6	8	18
i	0	20	4

Figure. A.3: Example 2.

each computer player is programmed to always make *the best choices* in terms of *the highest payoff* it could possibly make for itself *believing that all other players also make choices following with this rule*. It means that each of the three computer players will try to earn the highest payoff for itself *expecting that you and the other two players will also try to earn the highest payoffs for themselves expecting that everyone else in the game are also trying to earn their highest payoffs*. In other words, if you choose your own actions in the way as each computer player is programmed to choose theirs, you will earn the highest possible payoff *depending on the choices made by the other 3 players*. Recall that your payoff depends on your action and Player 2's action if you are Player 1 and so on.

We expect you to make choice within 2 minutes in each round. A timer will appear on the top on the screen. The timer will become red and flashing to remind you when there is 1 minute left. In addition, when there is only 30 seconds left, the text 'Please Make Your Choice!' will appear next to the timer. Though you will not get any penalty in terms of payoffs if you cannot make your decision within 2 minutes, participants who finish this part faster than other participants, will see a Waiting Page. Thus, please pay attention to the timer and do not keep other participants waiting. Finally, during this part, you will not receive any information either on other players' chosen actions or any players' payoffs including your own payoff until the end of the experiment.

Your Earnings

You will earn £5 for showing up for the experiment.

In addition, in this part, 1 out of 8 rounds will be randomly selected for payment at the end of the experiment. For example, suppose your payoffs in each of 8 rounds are 10 tokens, 16 tokens, 18 tokens, 12 tokens, 8 tokens, 4 tokens, 12 tokens, and 20 tokens respectively. Suppose the second round is selected for payment, so that your payoff for this part is 16 tokens. Your token earnings will be converted to pounds using exchange rate of £0.15/token. Thus, your payment is £2.40. Because you do not know which round will be selected until the end of the experiment, it is in your interest to do your best in each and every round in order to earn as much money as you can in this part.
Furthermore, you will have opportunities to earn more money in the further parts of this experiment. You will be informed of how you can earn money in each further part in separate instructions after you finish this part.
Now please click 'Next' button to move to comprehension quiz.

Instructions for Part 2

Please read the following instructions for Part 2 carefully. If you have any questions, please raise your hand, and an experimenter will come to help you in person.

Details

[Control: Irrelevant Tutorial] In this part of the experiment, you will be shown on screen the basics of colour addition and subtraction. You will observe how the two mechanisms work. You have to spend *at leas 5 minutes* on this part. The 'Next' button for you to move to next page will appear in 5 minutes. Moreover, you can spend *at most 10 minutes* on this part. The system will advance you to the next part automatically in

10 minutes. Finally, you will see a Waiting Page if you finish this part faster than other participants.

[Treatment: Relevant Tutorial] In this part of the experiment, you will be shown on screen how 4 computer players would make choices in the same tasks as in Part 1. You will observe which actions they would choose and which payoffs they would get. You have to spend *at least 5 minutes* on this part. The 'Next' button for you to move to next page will appear in 5 minutes. Moreover, you can spend *at most 10 minutes* on this part. The system will advance you to the next part automatically in 10 minutes. Finally, you will see a Waiting Page if you finish this part faster than other participants. **Note that computer players in this part are programmed in exactly the SAME way as in Part 1.**

Your Earnings

You will not be paid for this part. But in addition to the show-up fee and your earnings in Part 1, you will have a chance to earn money in the further parts of this experiment. You will be informed of how you can earn money in the next part in the separate instructions after you finish this part.

Please click 'Next' button to move to Part 2.

Instructions for Part 3

Please read the following instructions for Part 3 carefully. If you have any questions, please raise your hand, and an experimenter will come to help you in person.

Details

Similarly to what you have done in Part 1 of this experiment, in this part of the experiment, you will be asked to play 4-player games against 3 computer players again for 8 rounds. The payoff tables in this part will be different from those you have seen previously in Part 1. However, the procedure is the same as in Part 1, so you may wish to refer to the paper instructions for Part 1. Again, you will see a Waiting Page if you finish this part faster than other participants. **Note the computer players are programmed in exactly the SAME way as in Part 1.**

Your Earnings

Similarly to Part 1, one out of 8 rounds will be randomly selected for payment, and your earnings will be determined by your chosen action as well as the action of your counterpart Player in that randomly chosen round, and the exchange rate is £0.15/tokens. These earnings will be added up to your show-up fee and your earnings in Part 1. Furthermore, you will have a chance to earn money in the remaining two parts of this experiment, and you will be informed of how you can earn money in each part in separate instructions after you finish this part.

Please click 'Next' button to move to Part 3.

Instructions for Part 4

Please read the following instructions for Part 4 carefully. If you have any questions, please raise your hand, and an experimenter will come to help you in person.

Details

In this part, you will be asked to submit your prediction of the number of rounds that you think you have chosen the actions that would give you the highest possible payoff *depending on the computer players' choices* in each of Part 1 and Part 3. You will be informed of the actual number of rounds you have chosen the best-paying action in these two parts at the end of the experiment, so that you could compare your predictions to your actual performance. Finally, you will see a Waiting Page if you finish this part faster than other participants.

Your Earnings

Your predictions for both of these two parts will be paid. The prediction payoff for each part depends on the absolute distance of each predicted number relative to the actual number, and this will be done separately for each of the two parts. The closer is your prediction to the actual number, the higher payoff you will get (see the Table below).

The distance between your prediction and the actual number	Payoff
0	3.00 token
1	2.70 token
2	1.80 token
3	0.30 token
4	0.00 token
5	0.00 token
6	0.00 token
7	0.00 token
8	0.00 token

Table A.9: The payoffs for predictions in each of Part 1 and Part 3 separately.

For example, suppose that in Part 1 the number of rounds where you chose the best paying action is 5 - so that your actual number is 5. If your prediction is 5, then the distance is 0, so you will get 3.00 tokens for this prediction. And suppose that in Part 3 your actual number is 4, and your prediction is 3, then the distance is 1, so you will get 2.70 tokens for this prediction. Thus, your payoff in this Part is the sum of the payoffs to these two predictions which is 5.70 tokens. Again, the exchange rate is £0.15/token. Note that in the above example of a payment for prediction in Part 3 you could have also gotten 2.70 tokens if instead you predicted 5 - simply because the absolute distance between 4 and 5 is also 1. That is, what matters for your payment is how how high or low is your prediction, but how close your prediction is to the actual number.

These earnings will be added up to your show-up fee and your earnings in the previous parts. Furthermore, you will have opportunities to earn more money in final part of this experiment. You will be informed of how you can earn money in the next and final part in the separate instructions after you finish this part.

Please click 'Next' button to move to Part 4.

Instructions for Part 5

In this part, we ask you to answer a set of questions on the screen.

Your Earnings

We will pay you £3 for answering all questions in this part.

Thus, your total payment in this experiment is the sum of your token payoffs in Parts 1, 3, and 4 converted into pounds at the exchange rate of £0.15/token plus the show-up fee of £5 and £3 for answering all questions in part 5:

Final payment = (payoff from Part 1 + payoff from Part 3 + payoff from Part 4)×£0.15/token + £3 payoff from Part 5 + £5 show-up fee.

Please click 'Next' button to see the questions.

A.4 The Relevant Tutorial

In this part of the experiment, you will be shown on screen how 4 computer players, **which are programmed to behave in exactly the SAME way as in Part 1**, would make choices in the games played in Part 1. You will observe which actions they would choose and which payoffs they would get.

Recall that computer players in Part 1 were programmed to always consider all actions available to them, and select the action which would give them the highest payoff they can get – given that they believe other players are also doing so. And the game now is played by four computer players who make choices using this rule.

Example 1

Consider a game depicted in Figure 1 as played by four players who all are striving to get the highest payoffs and they know that all of their opponents are also striving to achieve their own payoffs as high as they can.

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f
a	**30**	6	20
b	4	20	**32**
c	12	**40**	6

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i
d	32	**24**	4
e	**36**	20	10
f	26	18	**16**

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	**18**	20	26
h	14	8	**36**
i	6	**24**	28

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	8	12	24
k	**20**	**16**	**32**
l	18	14	20

Figure. A.4: The highest payoff each player (whose actions are listed in the rows) can get when their opponent chooses a specific action (listed in the columns).

Let's start looking at this game from Player 1's perspective. What action would give this player the highest payoff – if its opponent chooses a specific action? Imagine that Player 2 chooses action d in Example 1 below. Then Player 1 would get the highest payoff by choosing a because action a gives it payoff of 30 tokens (which is highlighted in Figure 1), which is higher than the payoffs by choosing either action b or c in this scenario. Similarly, if Player 2 chooses action e, then action c will give Player 1 the highest payoff of 40 tokens; and if Player 2 chooses f, then action b will give Player 1 the highest payoff of 32 tokens (as highlighted in Figure 1).

Notice that Player 1's highest paying choice depends on Player 2's choices. Thus, in order to get the highest possible payoff Player 1 needs to figure out which action Player 2 would choose. But remember that Player 2 is also programmed to choose the highest paying choice for itself. So if you now look at the problem from Player 2's perspective, you would see that Player 2 would choose e, d, f when Player 3 chooses g, h, i respectively (again, the payoffs to the best paying actions are highlighted in Figure 1). But again, Player 2's highest paying choice depends on Player 3's choices.

So to predict Player 2's choice, we need to predict what Player 3 is going to choose. But, if Player 3 wants the highest payoff, this player would choose g, i, h when Player 4 chooses j, k, l respectively.

Interestingly, when we look at the problem from Player 4's perspective, it is clear that this player's highest paying action is k – for every action chosen by Player 1. That is, for Player 4, action k is "dominant", i.e. it is the best action Player 4 could take no matter what action Player 1 would choose. This means that Player 4 would always choose action k, and doesn't need to worry about Player 1's action.

An important feature in Example 1 is that it has Player 4 in a "key" position, as this player's choice starts the chain of "best responses" of all players. To observe how this chain works, see the highlighted action in Figure 2. Note that Player 4 always chooses action k no matter what action Player 1 would choose. Then, Player 3, whose opponent is the "key" Player 4, will choose action i expecting that Player 4 will choose action k – because Player 3 can see that action k is the best choice Player 4 can make if Player 4 wants the highest payoff. Meanwhile, Player 3 also wants to have a payoff as high as possible, so it will choose action i which gives itself the highest payoff when Player 4 chooses action k.

Similarly, Player 2 expects that Player 4 would choose k leading Player 3 to choose action i, and so Player 2 will choose action f. Finally, Player 1 chooses action b such that it will earn the highest payoff expecting that Player 4 will choose k, Player 3 will choose action i and Player 2 will choose action f. Therefore, these four players settle down on actions b, f, i and k respectively, and get the highest payoffs given that all other players are also striving to get the highest payoffs.

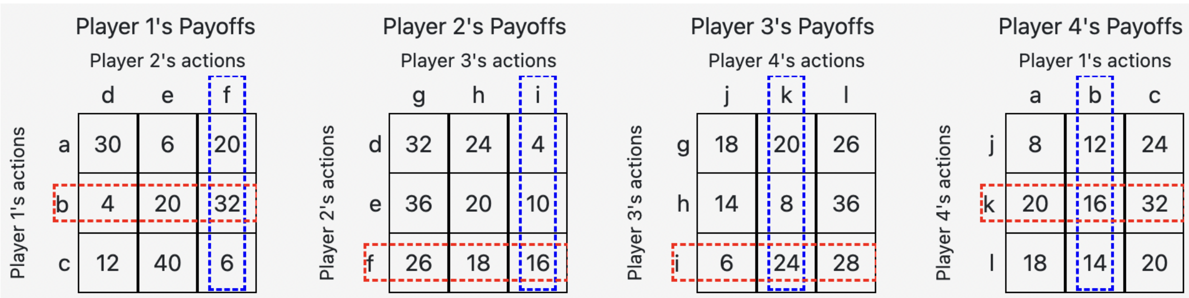


Figure. A.5: The chain of "best responses" of all players.

Example 2

Now consider Example 2 in Figure 3. Recall that every computer player is programmed to choose an action which would give itself the highest payoff given that all other players are doing so, and the game is played by four such computer players.

Similarly to what we have done for Example 1, let's identify which action would give each player the highest payoff when their opponent chooses a specific action. If Player 2 chooses action d, action a would give Player 1 the highest payoff of 30 tokens – as actions b and c only give payoff of 4 and 12 tokens, both lower than 30 tokens (see the highlighted payoff in Figure 3). Similarly, when Player 2 chooses action e or action f, in order to maximize its payoff Player 1 would choose action c or b, respectively.

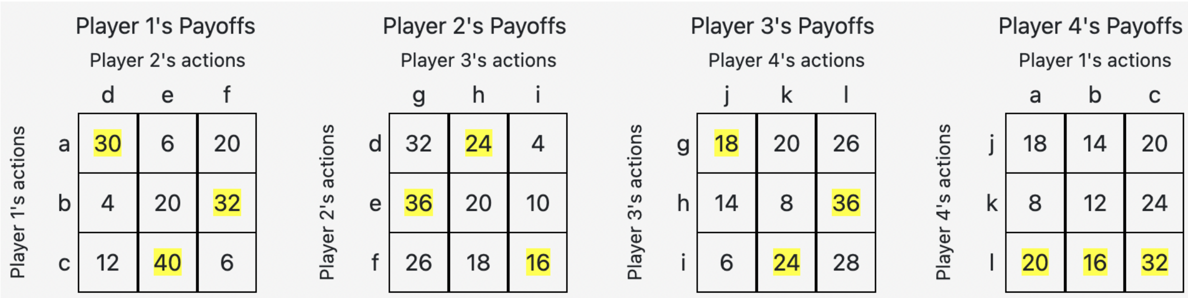


Figure. A.6: The highest payoff each player (whose actions are listed in the rows) can get when their opponent chooses a specific action (listed in the columns).

Likewise, Player 2 would choose actions e, d and f when Player 3 chooses action g, h, i, respectively; Player 3 would choose actions g, i, and h when Player 4 chooses actions j, k, l, respectively. And Player 4 would always choose action l regardless of actions chosen by Player 1 (see the highlighted payoffs in Figure 3).

That is, just like in Example 1, Player 4 is in a "key" position in Example 2, and has a "dominant" action l (see the highlighted action in Figure 4). Now, Player 3, who faces the "key" Player 4, will choose action h expecting that Player 4 will choose their "dominant" action l. Since Player 3 also want to make the highest payoff, it will choose action h as it knows that Player 4 will choose l for sure. Similarly, Player 2 expects that Player 3 will choose h expecting action l to be chosen by Player 4, it will choose action d. At the same time, Player 1 expects that Player 2 will choose d, Player 3 will choose h, and Player 4 will choose l, so Player 1 will choose action a to maximize payoff. Therefore, every player settles down at its best choice and earn the highest payoff using this rule.

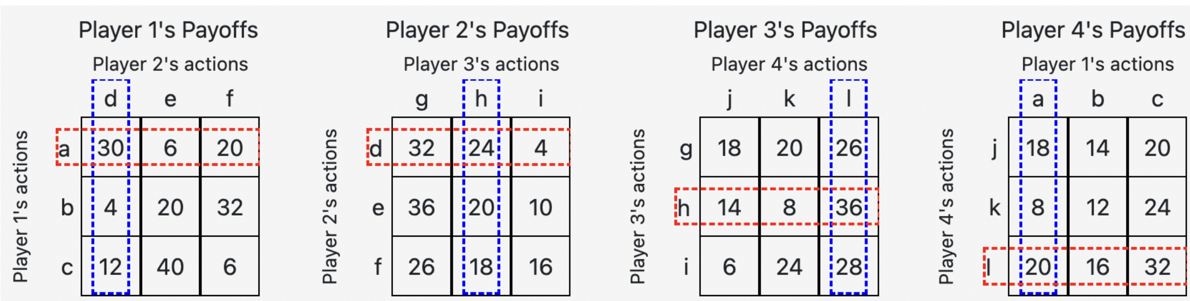


Figure. A.7: The chain of "best responses" of all players.

A.5 The Irrelevant Tutorial

In this part of the experiment, you will be shown on screen the basics of colour addition and subtraction.

The colour of light is the result of human vision perception which is triggered by the associated wavelength of the light when the light reaches human eyes.

Scientists developed a simple method to represent, produce and implement as many as possible different colours. This method involves two different mechanisms - colour addition and colour subtraction.

Colour Addition

Broadly speaking, colour addition is a mechanism of perception of colours of bright, light-emitting, objects. To understand colour perception, we need to understand some basics of the biology of the human eye. The human eye is a complex organ. The eye's "retina" (which is the coating of the interior surface at the back of the eye) contains special light sensitive cells called "cones" as they are sensitive to different wavelengths. Some of these cone cells are activated by light of long wavelengths (so called L cone cell), some are activated by light of medium wavelengths (M cone cell), and others are activated by light of short wavelengths (S cone cell). Thus, different types of cone cells are sensitive to different colours.

Each person's eye normally has all three of these different types of cones cells. Importantly, when more than one of cones is stimulated, then the stimulation from those activated cones is combined, or "mixed", and one perceives a colour which is different from the colour if only either one of the cones is simulated. This mechanism of mixing additive colours (which is akin to mixing stimulations of different types of eye cones) is called "colour addition". This mechanism simply implies that if one combines, or "mixes", two additive colours, one gets a different, third, colour. In order to represent as many colours as possible, three primary colours are chosen as the base colours for mixing in colour addition - red, green and blue. Thus, a specific, non-primary colour could be obtained by mixing primary colours using different intensities (where colour "intensity" means the degree of colour brightness). This means that the final colour you perceive is the result of the combination of wavelengths of primary colours (which is due to the different ways these wavelengths activate the cone cells).

The colour addition as the mechanism for colour perception is widely used in the modern world to create electronic displays, in colour television and in theatre stage lighting. For example, your computer monitor uses different intensities of the three primary colours, – red, green and blue – which your eyes then perceive as different colours.

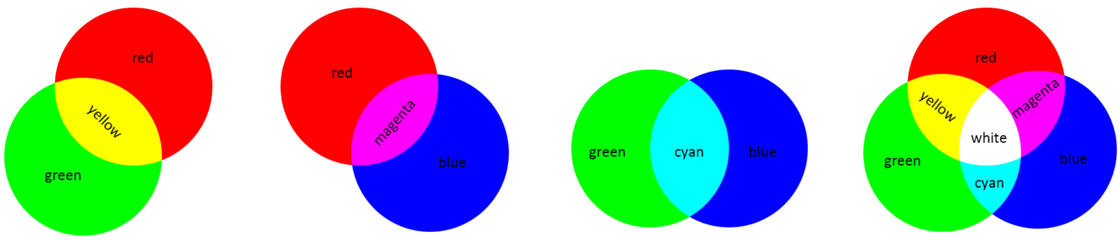


Figure. A.8: Examples of colour addition.

Figure A.8 is an example of the addition of these three primary colours. The leftmost panel shows that yellow could be obtained by adding green and red each of which has the full intensity (i.e., the red and green lights are turned on as bright as possible). Similarly, adding red and blue using full intensity gives magenta, and adding green and blue using full intensity gives cyan. The rightmost panel demonstrates that when the three primary colours are added together with full intensity, we will obtain white. Note that if you perceive white, it must be some combination of other colour lights since there is no visible single wavelength corresponds to white light. Furthermore, when the computer screen does not emit light, we perceive the colour of the blank screen as dark or black; when the screen emits light and if we perceive white light, this must be the result of the combination of red, green and blue lights of the screen.

Colour Subtraction

The other mechanism, colour subtraction, refers to the mixing of subtractive colours. Broadly speaking, it is the mechanism of how we perceive the colours of objects that are not illuminated or opaque. It is used in printing where the primary colours typically are cyan, magenta and yellow. You might also have experience of creating new colours by subtraction in your early childhood - when you mixed paints to create your pictures in your painting classes.

When light hits the surface of an object, the atoms which covers that surface work as a light "filter", absorbing some light wavelengths, and reflecting other wavelengths. Recall that from colour addition, we have known that white light comes from the combination of wavelengths of the primary colours - red, green and blue. When white light passes through a filter which absorbs some colour(s) from the white light, the filter reflects the remaining light colours which reach our eye, generating the final colour we perceive. So the generated colour (which is the colour of the filter) is simply what is left from the combination of the three primary colours which comprise the white light after some of these primary colours get absorbed, or "subtracted". So one can explain the final colour that we perceive using subtraction of different intensities of three primary colours from white, as the difference of wavelength between white and primary colours.

Figure A.9 is an example of subtraction of primary colours. Consider the leftmost panel, which shows that when one mixes magenta and yellow, one gets red. To understand it, first note that magenta is the result of absorption of green from white: $white - green = red + blue = magenta$. Similarly, yellow is the result of absorbtion of blue from white:

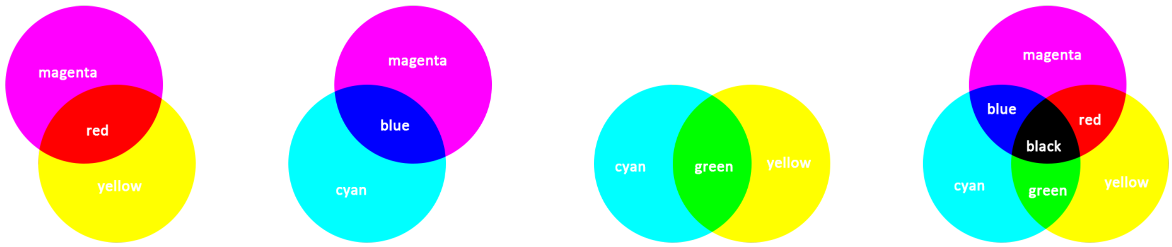


Figure. A.9: Examples of colour subtraction.

$white - blue = red + green = yellow$. Now observe the intersection of these two subtractive colours. Because magenta is what is left of white light after its green component is removed (subtracted), and yellow is what is left of white light after its blue component is subtracted, when we mix magenta and yellow, we get red as white light with both green and blue components removed, or $white - green - blue = red$.

Similarly, cyan is what is left of white light after its red component is removed, or $white - red = green + blue = cyan$. Thus, when we combine magenta and cyan (as in the second panel from the left in Figure A.9), we get what is left of white light after both green and red are removed, which is blue: $white - green - red = blue$.

Furthermore, when we combine cyan and yellow (as in the third panel from the right in Figure A.9), we get what is left of white light after both red (because of cyan) and blue (because of yellow) are removed, which is green: $white - red - blue = green$.

Finally, as the rightmost panel in Figure A.9 shows, when we combine all three - cyan, magenta, and yellow, we get what is left after red, green, and blue are removed (respectively). But that means that we removed all three primary colours from the white light - so now surprise we are left with no colour - which means we are left with black: $white - red - green - blue = black$.

Appendix B for Chapter 2

B.1 Additional Tables

Table B.1: Abilities and levels assigning subjects to the higher type when they can be classified into two types with an error.

		Against Others						
		L0	L1	L2	L3	L4	Total	
Against Yourself	A0	103	18	5	0	0	126	70.39%
	A1	8	28	1	0	0	37	20.67%
	A2	3	4	2	0	0	9	5.03%
	A3	1	1	1	0	0	3	1.68%
	A4	1	0	0	0	3	4	2.23%
Total		116	51	9	0	3	179	
		64.80%	28.49%	5.03%	0%	1.68%		

Notes: Percentages in the rightmost column and the bottom row. Counts in the table cells. The shades of the cells represent the changes in reasoning depth: light - $Ak > Lk$; medium - $Ak < Lk$; dark - $Ak = Lk$.

Table B.2: Abilities and levels (two errors).

		Against Others					Total	
		L0	L1	L2	L3	L4		
Against Yourself	A0	66	18	9	1	0	94	52.51%
	A1	7	47	2	0	0	56	31.28%
	A2	7	4	5	1	0	17	9.50%
	A3	0	1	3	3	0	7	3.91%
	A4	1	1	1	0	2	5	2.79%
Total		81	71	20	5	2	179	
		45.25%	39.66%	11.17%	2.79%	1.12%		

Notes: Percentages in the rightmost column and the bottom row. Counts in the table cells. The shades of the cells represent the changes in reasoning depth: light - $Ak > Lk$; medium - $Ak < Lk$; dark - $Ak = Lk$.

Table B.3: 2-type finite mixture model.

	Type 1	Type 2
Time spent on games (sec)		
CRT	5.820***	19.472**
	(2.781)	(8.099)
Age	0.323	-0.038
	(0.410)	(1.561)
Study Level	-4.716	-69.986**
	(7.284)	(34.246)
Female	-2.095	-4.239
	(13.639)	(29.994)
Log Hits	2.800	7.012
	(1.747)	(4.388)
Science & Economics	-9.001	42.458
	(10.695)	(30.130)
Risk	-7.958**	-6.025
	(3.501)	(7.391)
Constant	111.243**	395.054
	(45.608)	(248.429)
Mean of Time Spent	83.592***	202.968***
	(5.732)	(20.285)

Notes: $N = 143$ due to a few subjects not reporting gender and major. Dependent variable: time spent on games. Robust standard error. Type 1 subjects on average spent 84 seconds on the games while Type 2 subjects spent about 203 seconds.

Table B.4: Abilities and levels (one error) after screening.

		Against Others					Total	
		L0	L1	L2	L3	L4		
Against Yourself	A0	22	7	1	0	0	30	58.82%
	A1	2	11	0	0	0	13	25.49%
	A2	0	2	1	0	0	3	5.88%
	A3	0	1	1	1	0	3	5.88%
	A4	1	0	0	0	1	2	3.92%
Total		25	21	3	1	1	51	
		49.02%	41.18%	5.88%	1.96%	1.96%		

Notes: Percentages in the rightmost column and the bottom row. Counts in the table cells. The shades of the cells represent the changes in reasoning depth: light - $Ak > Lk$; medium - $Ak = Lk$; dark - $Ak < Lk$. Subjects with Type-2 probability ≤ 0.5 based on model in Table B.3 excluded. That is, subjects who spent too little time on the games are screened out. Mean $Ak = 0.71$ of the Mturk sample, significantly lower than mean $Ak = 1.02$ of the student sample in Chapter 1 (one-size t-test $p = 0.044$).

Table B.5: Summary Statistics. Mean/Share (SD).

	Dropouts	Finished
Female	0.324 (0.470)	0.371 (0.484)
Native	0.993 (0.084)	0.989 (0.105)
Age	37.27 (10.15)	37.63 (11.47)
Risk	7.182 (2.488)	6.190*** (2.679)
Study Level	3.811 (0.804)	3.821 (0.869)
Science & Economics	0.432 (0.497)	0.531* (0.500)
CRT	1.439 (1.751)	2.743 (2.403)
Hits	42408.3 (258701.5)	28868.0*** (152487.8)
N	148	179

Notes: *Female* is equal to 1 if a subject is female, 0 if a subject is male, and missing if gender is not reported. *Native* takes a value of 1 if a subject is an English native speaker. *Risk* takes a value from 0 (risk aversion) to 10 (risk loving). *Study level* takes takes a value from 1 to 5 when a subject has not finished high school, finished high school, went to some college/university without completed degree, got undergraduate degree, got postgraduate degree respectively. *Science & Economics* take a value of 1 if a subject studied or is studying natural science, applied science, and economics. See Frederick (2005); Toplak et al. (2011) for CRT questions. *Hits* is the number of assignments a subject has completed for Mturk. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ according to Mann-Whitney test of difference of each variable between dropout and finished group.

B.2 Instructions

In this part of the study, you will earn tokens by playing a four-player game. At the end of the study, the tokens you earn will be converted into **bonus** as will be described below. The greater are your token earnings, the greater are your bonus.

The Rules

Each of four players will choose one out of four available actions. Each player's payoff (in tokens) will depend directly on the combination of own action choice and the action choice of one other player. Each player's payoff possibilities will be represented in that player's payoff table, similar to the one in Figure B.1 below.

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f
a	22	6	20
b	36	10	2
c	6	28	14

Player 2's payoffs

Player 2's actions \ Player 3's actions	g	h	i
d	20	12	16
e	26	8	14
f	4	32	12

Player 3's payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	26	16	6
h	2	28	18
i	10	8	30

Player 4's payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	14	10	24
k	2	28	18
l	26	6	16

Figure. B.1: An example of a 4-player game.

Suppose you are Player 1 in this example. Your payoff possibilities are listed in the first table (as indicated in bold font above that table).

- As Player 1, you can only choose among 3 actions, *a*, *b*, or *c*; while Player 2 can only choose among 3 actions,*d*, *e*, or *f*.
- As Player 1, your action will determine the row of your payoff table, and Player 2's action will determine the column of the table.

- Your (Player 1's) token payoff is determined by the number in the cell corresponding to the combination of the actions chosen both by you (as Player 1) **and** by Player 2 ***directly***.

Payoffs of Player 2, Player 3, and Player 4 are listed in the other three tables (as denoted in bold).

- Player 2 can choose action *d*, or *e*, *f*. and this player's payoff depends upon the action (s)he chooses and the action Player 3 chooses.
- Player 3 can choose action *g*, *h*, or *i*, and this player's payoff depends upon the action (s)he chooses and the action Player 4 chooses.
- Player 4 can choose action *j*, *k* or *l*, and this player's payoff depends upon the action (s)he chooses and the action which you, as Player 1, choose.

Example: suppose Player 1 chooses *b*, Player 2 chooses *f*, Player 3 chooses *g*, Player 4 chooses *l*. In Figure B.2 below, the corresponding row and column of the chosen actions are highlighted (in blue); the corresponding token payoff of the player whose name in bold (at the top of the table) is in the cell at the intersection of the highlighted row and column (in red). Thus, Player 1 earns 2 tokens, Player 2 earns 4 tokens, Player 3 earns 6 tokens and Player 4 earns 6 tokens.

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f
a	22	6	20
b	36	10	2
c	6	28	14

Player 2's payoffs

Player 2's actions \ Player 3's actions	g	h	i
d	20	12	16
e	26	8	14
f	4	32	12

Player 3's payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	26	16	6
h	2	28	18
i	10	8	30

Player 4's payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	14	10	24
k	2	28	18
l	26	6	16

Figure. B.2: The above example, with highlighting.

You can move your mouse over a cell of a table to see the corresponding actions of two players in the payoff tables. Also you can highlight a particular cell in yellow by clicking it. (To undo highlighting, click the cell again.)

Your Decisions and Tasks

There will be **two rounds**. In each round you will see a set of payoff tables similar to the above example in Figure B.1. The two rounds will have different sets of payoff tables (and both sets of payoff tables will be different from those depicted in Figures B.1 and B.2).

In each round you will face two types of tasks, called Task O and Task Y, presented to you simultaneously on the same page, as depicted in Figure B.3. In each round and each task, you will be asked to make 4 decisions, by playing in **each** of four player roles.

That is, in each round, you will make 4 Task O decisions - in the role of Player 1, in the role of Player 2, and so on. You will make a total of 8 Task O decisions, labelled as O1-O4 in round 1 and O5-O8 in round 2. Same for 8 Task Y decisions you need to make, labelled as Y1-Y4 in round 1 and Y5-Y8 in round 2.

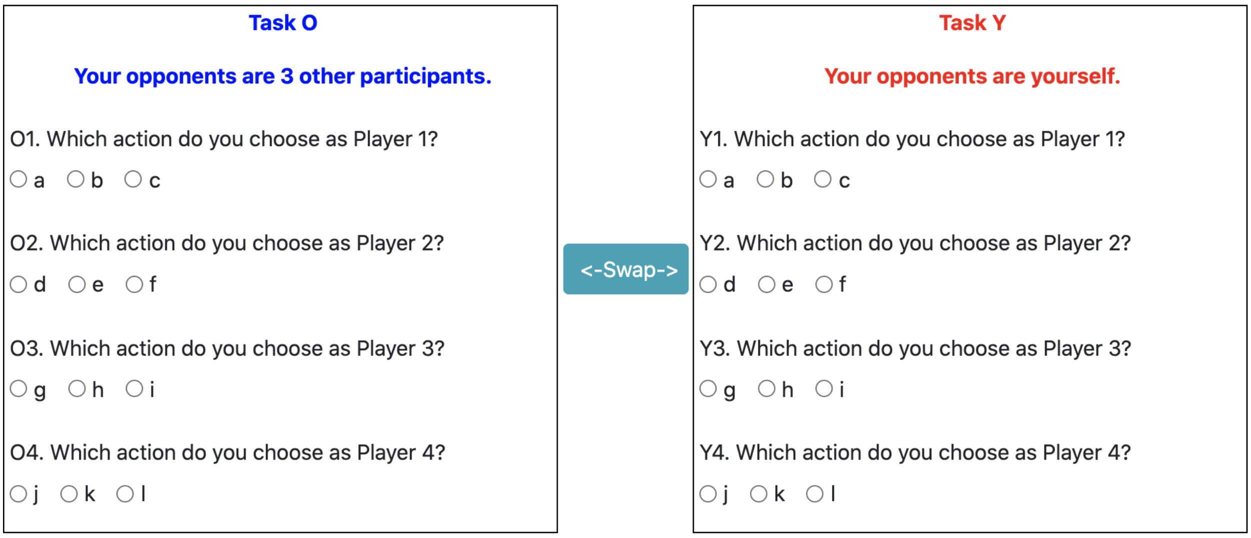


Figure. B.3: An example of the two tasks (in round 1).

Task O: 'Play Against Others'

- In Task O, each of your three opponents will be ***other Mturkers*** who also are participating in this study. The other participants face exactly the same payoff tables as you do, and exactly the same tasks.

- You and the other participants in the study will remain mutually anonymous to each other. Thus, in Task O you ***do not know which actions other players choose***.

- After you make all of your decisions, you will be matched in a four-player group with three other participants, and your action choices will be matched against the corresponding actions chosen by these three other participants. That is, the action you chose in the role of Player 1 will be matched with Player 2's action chosen by one participant, Player 3's action chosen by another participant, and Player 4's action chosen by a third participant, and the token payoffs of the participants in your group will be determined according to the rules of the game described above. Similarly, your action choice as Player 2 will be matched with the action choices by three other participants as Player 1, 3, and 4, and so on. This procedure results in you and these three other participants ***taking turns*** at playing in each of four player roles.

  – Note that it is possible that due to participant numbers, some groups might have fewer than four participants. If such (infrequent) situation does happen, in order to insure that indeed all payoffs in Task O are determined by real participants, the corresponding actions of randomly selected study participants will be 'reselected' to complete groups which do not have enough players, – without altering the payoffs of those 'reselected' participants. As you and your opponents remain anonymous to each other, this procedure of re-selection of actions into incomplete groups should not affect your decision making.

Task Y: 'Play Against Yourself'

- In task Y, each of your three opponents is **yourself**. That is, your action choice made in the role of Player 1 will be matched against your action choice in the roles of Player 2, Player 3, and Player 4. That is, all four decisions in each round will determine your and only your payoff according to the rules of the game described above.

- Thus, in Task Y you ***know which actions other players choose***.

You can switch the position of two tasks on the screen as you wish by clicking 'Swap' button. You can choose same or different actions across the two tasks. **It is in your best interest to choose an action which you believe would give you the most tokens (bonus).**

Your Earnings (Reward+Bonus)

At the end of the study, **1 decision** will be randomly selected for **bonus**.

- First, either Task O or Task Y will be randomly selected, with equal probability of 1/2.

- Second, among your 8 decisions for the selected Task, only one decision will be randomly selected, with equal probability of 1/8.

- Example: suppose Task Y is randomly selected, and the selected decision is Y7. That means that your bonus will be determined by your token payoff in your role of Player 3 in round 2 of Task Y. Suppose further that in round 2 of Task Y the payoff tables that you faced and the decisions that you made in each player role are the same as depicted in Figure B.2. Then you earn 6 tokens (as Player 3).

- Each token earned is worth \$0.10 (or 10 tokens = \$1.00).

If you complete all parts of the study, your final earnings will be **the reward of \$2.50 plus the guaranteed bonus earned in the four-player game** with range of \$0.20-\$3.60 as advertised.

B.3 Screenshots of Games

Round 1/2

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f
a	14	6	16
b	26	22	18
c	18	4	14

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i
d	12	24	16
e	18	14	28
f	26	8	18

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	26	14	6
h	8	10	28
i	12	30	14

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	32	4	6
k	6	2	36
l	4	26	12

Task O

Your opponents are 3 other participants.

O1. Which action do you choose as Player 1?
○ a ○ b ○ c

O2. Which action do you choose as Player 2?
○ d ○ e ○ f

O3. Which action do you choose as Player 3?
○ g ○ h ○ i

O4. Which action do you choose as Player 4?
○ j ○ k ○ l

<-Swap->

Task Y

Your opponents are yourself.

Y1. Which action do you choose as Player 1?
○ a ○ b ○ c

Y2. Which action do you choose as Player 2?
○ d ○ e ○ f

Y3. Which action do you choose as Player 3?
○ g ○ h ○ i

Y4. Which action do you choose as Player 4?
○ j ○ k ○ l

Next

Instructions

Round 2/2

Player 1's Payoffs

Player 1's actions \ Player 2's actions	d	e	f
a	26	22	18
b	18	4	14
c	14	6	16

Player 2's Payoffs

Player 2's actions \ Player 3's actions	g	h	i
d	12	24	16
e	18	14	28
f	26	8	18

Player 3's Payoffs

Player 3's actions \ Player 4's actions	j	k	l
g	26	14	6
h	8	10	28
i	12	30	14

Player 4's Payoffs

Player 4's actions \ Player 1's actions	a	b	c
j	32	4	6
k	6	2	36
l	4	26	12

Task O

Your opponents are 3 other participants.

O5. Which action do you choose as Player 1?
○ a ○ b ○ c

O6. Which action do you choose as Player 2?
○ d ○ e ○ f

O7. Which action do you choose as Player 3?
○ g ○ h ○ i

O8. Which action do you choose as Player 4?
○ j ○ k ○ l

<-Swap->

Task Y

Your opponents are yourself.

Y5. Which action do you choose as Player 1?
○ a ○ b ○ c

Y6. Which action do you choose as Player 2?
○ d ○ e ○ f

Y7. Which action do you choose as Player 3?
○ g ○ h ○ i

Y8. Which action do you choose as Player 4?
○ j ○ k ○ l

Next

Instructions

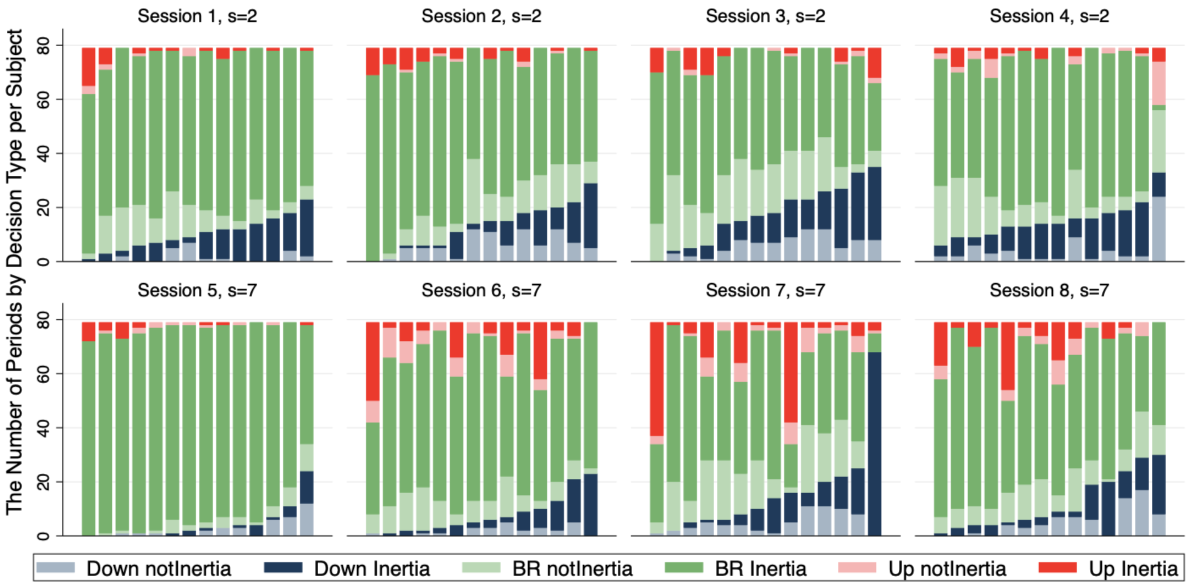
Appendix C for Chapter 3

C.1 Does the Inertia Cause Downward Deviations

In the main text section 3.3.4, we investigate the subjects' perceptions of signals and their responsiveness to signals with two sampling sizes as a mechanism driving the downward deviations. Here we study another possible mechanism - whether subjects are inertial while they should have caught up with changes in the signals. Specifically, if subjects were inertial, they would choose the same action regardless of the changes in the signals.

Let $inertia_{it}$ take the value of 1 if $Action_{it} = Action_{i,t-1}$, and 0 otherwise. We describe an action in period t by subject i as an inertial choice if $ineria_{it} = 1$. Figure C.1 depicts the number of periods that a subject played mBR, downward and upward deviations, grouped by $inertia_{it}$, where each bar corresponds to a subject. Among the downward deviations, there is a noticeable proportion of inertial choices, indicating that downward deviations might be caused by inertia.

We further test the inertia hypothesis by logit regressions of $inertia_{it}$. The explanatory variables include $s7_i$, $OwnSignal_{it}$ which is the fraction of the action chosen in period $t-1$ by subject i in the signal at the beginning of period t, $DSignal_{it}$ which is the dummy variable taking the value of 1 if the fraction of the action chosen in period $t-1$ in a signal decreases in period t. We say a signal in period t is good if $DSignal_{it} = 1$ in the sense that this signal encourages subjects to use the action chosen in the last period because $DSignal_{it} = 1$ indicates more players are playing the previously chosen action,



Notes: Sorted by downward deviations. Each bar corresponds a subject in a session. A choice is labelled by 'Inertia' if $Action_{it} = Action_{i,t-1}$, 'notInertia' otherwise. BR = myopic best response to signals. Down = downward deviation. Up = upward deviation.

Figure. C.1: The number of periods of each type of behaviors, per subject.

and thus their previous action choices are 'correct'; otherwise, a signal is bad. We also use interaction term $DSignal_{it} \times OwnSignal_{it}$ to explore whether there exist asymmetric effect of good and bad signals: it is reasonable to expect that when a signal is good, an increase in $OwnSignal_{it}$ would lead to larger increase in probability of taking the same action in two periods, i.e., $inertia_{it} = 1$, compared to the case in which a signal is bad. Model (1) in Table C.1 reports the basic estimates and model (2) includes the same set of controls as in the main paper. The standard errors are clustered at the session level.

Both models (1) and (2) sugget that subjects adjust their action choices to changes in signals. First, we find a significant negative effect of $DSignal_{it}$ on the probability of playing the same action, indicating that subjects would like to switch away from the previous action if the signal in the current period is bad. Second, we find a significant positive effect of $OwnSignal_{it}$ on the probability of playing the same action, indicating that subjects increase the probability playing the same action as in the last period if they observe more other subjects playing the action. In addition, the size of the effect of $OwnSignal_{it}$ is significantly smaller when signals are bad, i.e., the interaction term $DSignal_{it} \times OwnSignal_{it}$ has a significantly negative effect. That is, the effects of good and bad signals are asymmetric: subjects overreact to bad signals than to good signals. Such "win–stay, lose–switch" behavior in fact indicates that subjects closely monitored the signals, and adjusted their choices accordingly, rather than stayed inertial and chose the same action regardless of the changes in the signals.

Since two out of all sessions transit to the medium-efficient equilibrium e_2 instead of the most efficient equilibrium e_3, we further look at whether subjects are on average more reluctant to switch away from action 2 than from other actions. Building on models (1) and (2) in Table C.1, models (3) and (4) further include action dummies in period $t-1$ and their interaction terms with $OwnSignal_{it}$. Compared to action 2, choosing action 3 in period $t-1$ does not significantly change the probability of choosing itself again in period t; while choosing action 1 in period $t-1$ significantly decreases the probability (of choosing itself in period t). Thus, the switch rates away from the previous action choices are the same when the previous action choices are action 2 and 3. However, if the previous action choice is action 1, subjects switch away from it more often compared to the case where the previous action is action 2. The interaction terms of action dummies and $OwnSignal_{it}$ are not significant, i.e., the effects of $OwnSignal_{it}$ on the probability of choosing the same action as in the last period are the same for all the three action choices. An interpretation could be that subjects do not consider action 3 a riskier option than action 2 so that require no significantly larger increase in $OwnSignal_{it}$ to play the same action. Overall, we do not find evidence of inertia when the effect of signals is controlled for.

Result C.1. *Subjects do not have tendency to play the same action over time when the effect of signals is controlled, and thus downward deviations are not caused by inertia.*

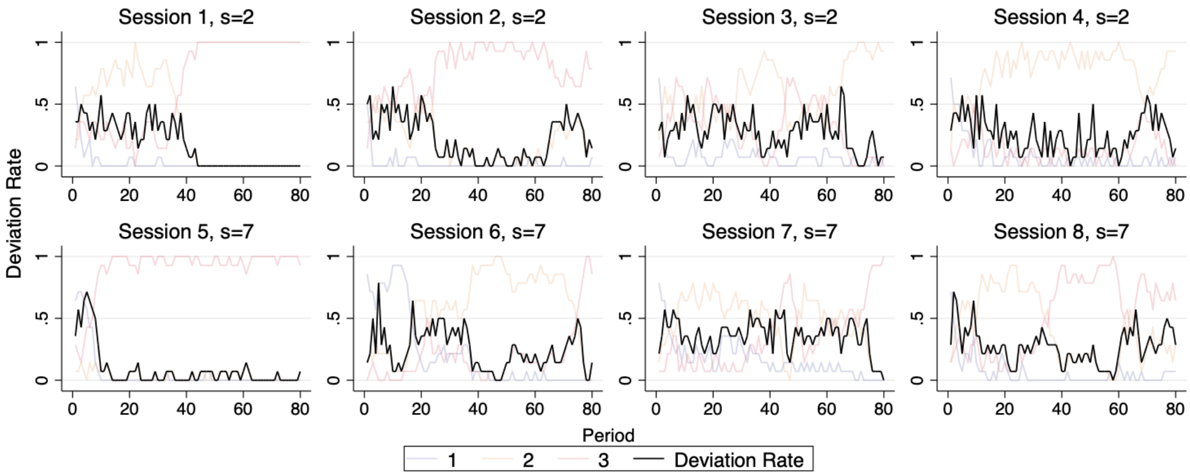
On the other hand, subjects on average switch away from action 1 *more often than from action* 2*, but there is no difference in the rates of switching away from action* 2 *and* 3*.*

Table C.1: Panel logit regressions. Test of Inertia.

$Inertia_{it}$	(1)	(2)	(3)	(4)
$s7_i$	0.235**	0.305***	0.312***	0.361***
	(0.108)	(0.090)	(0.116)	(0.112)
$DSignal_{it}$	-0.361**	-0.356**	-0.423***	-0.418***
	(0.144)	(0.145)	(0.135)	(0.140)
$OwnSignal_{it}$	3.763***	3.727***	3.562***	3.542***
	(0.244)	(0.235)	(0.226)	(0.226)
$DSignal_{it} \times OwnSignal_{it}$	-0.701**	-0.737**	-0.562*	-0.600**
	(0.304)	(0.308)	(0.296)	(0.299)
$Action1_{i,t-1} \times OwnSignal_{it}$			-0.842*	-0.864
			(0.469)	(0.526)
$Action3_{i,t-1} \times OwnSignal_{it}$			0.293	0.265
			(0.495)	(0.508)
$Action1_{i,t-1}$			-0.436**	-0.443**
			(0.184)	(0.206)
$Action3_{i,t-1}$			0.035	0.036
			(0.195)	(0.195)
$StudyLevel_i$		0.217		0.276
		(0.458)		(0.492)
$Female_i$		-0.381		-0.355
		(0.321)		(0.347)
$Economics_i$		0.105		0.072
		(0.134)		(0.152)
$nCRT_i$		0.199**		0.200**
		(0.084)		(0.087)
Age_i		-0.004		-0.010
		(0.074)		(0.078)
Constant	-0.289***	-1.214**	-0.210	-1.191*
	(0.092)	(0.576)	(0.146)	(0.622)
N	8848	8532	8848	8532

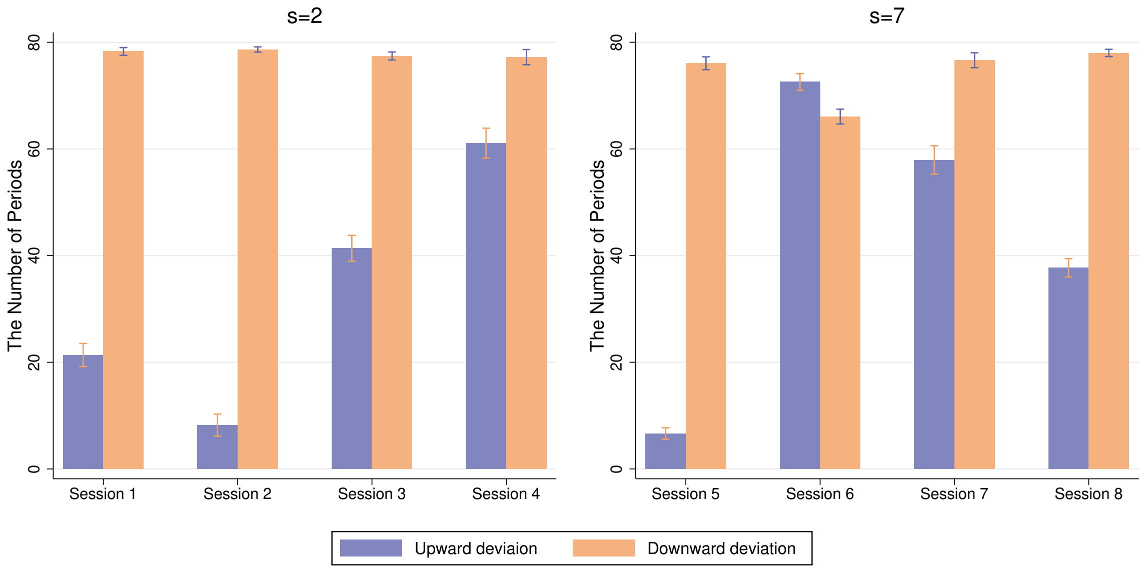
Notes: Session clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Fewer observations in columns (2) and (4) are due to missing values in $Female_i$ and Age_i.

C.2 Additional Tables and Figures



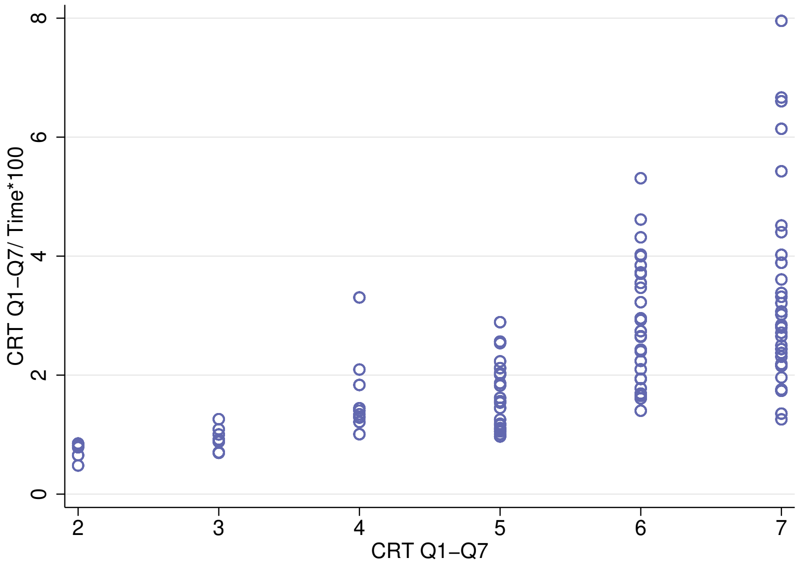
Notes: Transitions plotted at the background. In general, the deviation rate is high at the beginning of the game, then decreases to a low level as the population reached an equilibrium (or the neighborhood of an equilibrium). In the sessions transiting to e_2 and e_3 in sequence, the rise in the deviation rate is evident again when the population starts to transit to e_3 followed by another decrease. Thus, the deviation rate usually has a *hump shape* from the beginning to the end of the transitions.

Figure. C.2: The deviation rate over time by session.



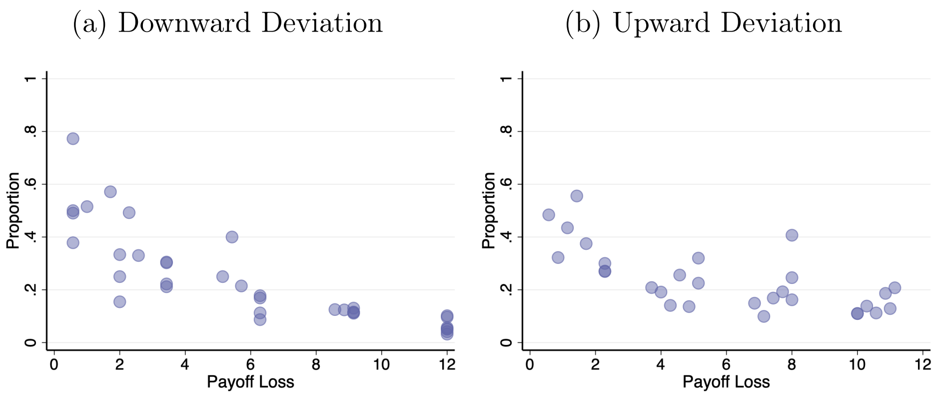
Notes: Error bars represent standard deviations.

Figure. C.3: The number of periods when two types of deviations are available.



Notes: There is significant positive correlation between the two measures using CRT: The Pearson's correlation coefficient is 0.6167***, and the Spearman's ρ is 0.6981***.

Figure. C.4: Relationship between raw CRT and normalized CRT.



Notes: Both upward and downward deviations are negatively correlated with the associated payoff loss: The Spearman's ρ is -0.9042*** for downward deviation and -0.6882*** for upward deviation. Furthermore, Upward deviation is significantly less correlated with the payoff loss than downward deviation ($\Delta\rho = 0.2160^{*}$).

Figure. C.5: Scatter plots of deviation rate and associated payoff loss given signals (full data set).

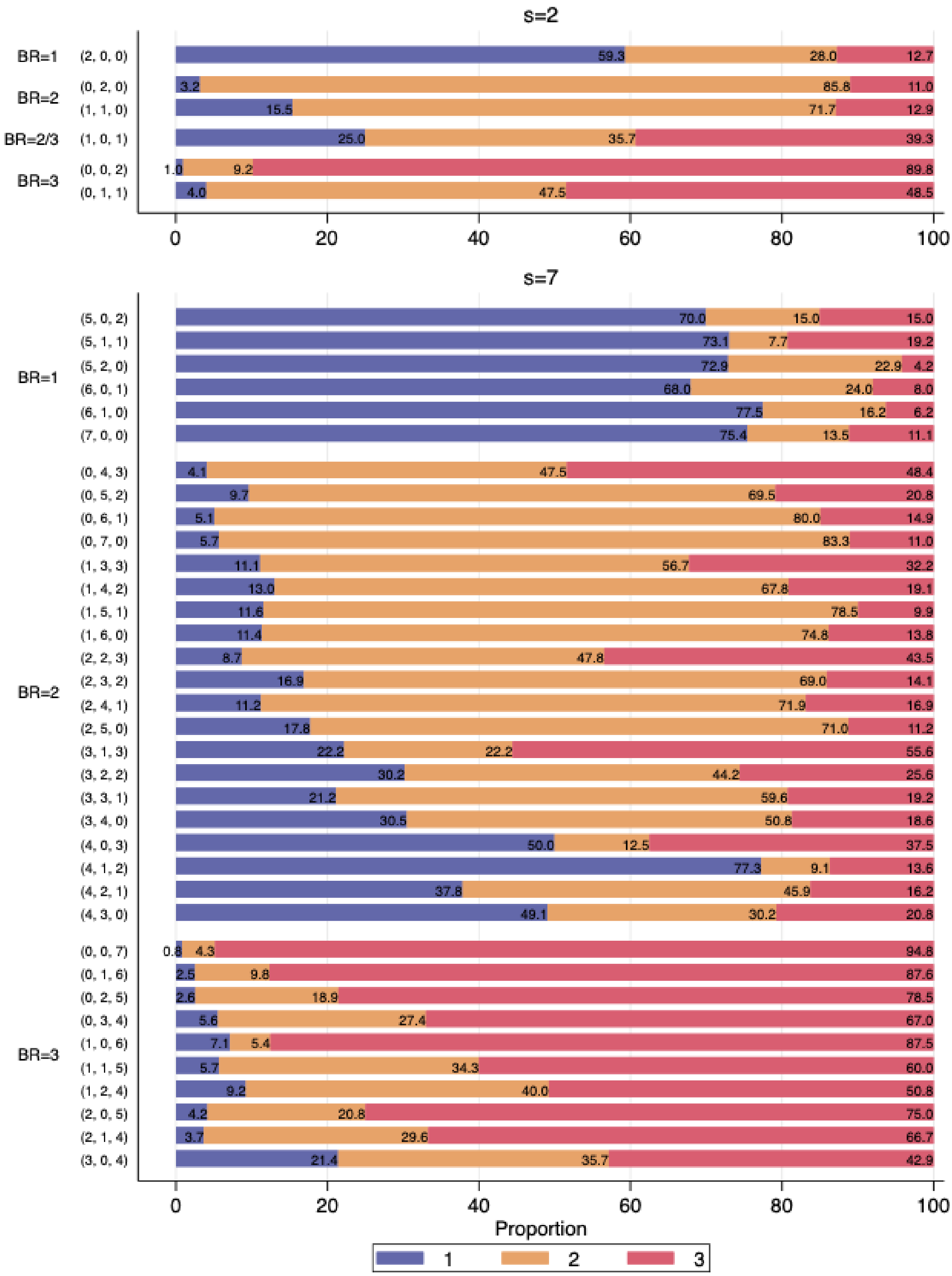
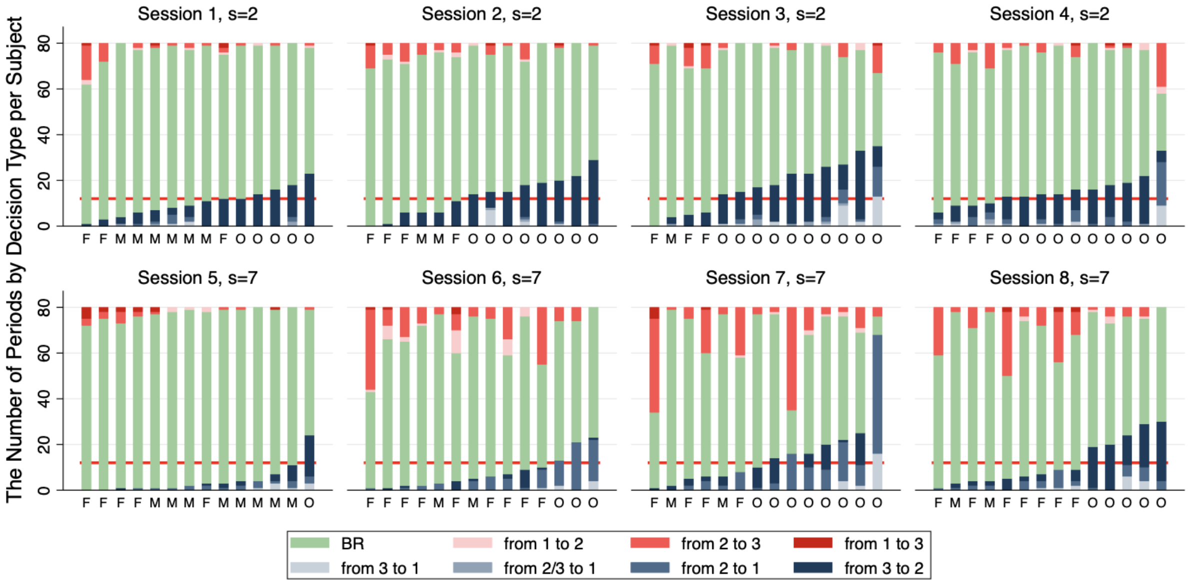


Figure. C.6: The proportion of action choices given each signal.



Notes: Sorted by downward deviations. The red horizontal reference line is 15% of time on downward deviations. Each subject is labeled by her type: F = FL player, M = Myopic player, O = Others.

Figure. C.7: The number of periods of BR, downward and upward deviations played by every subject.

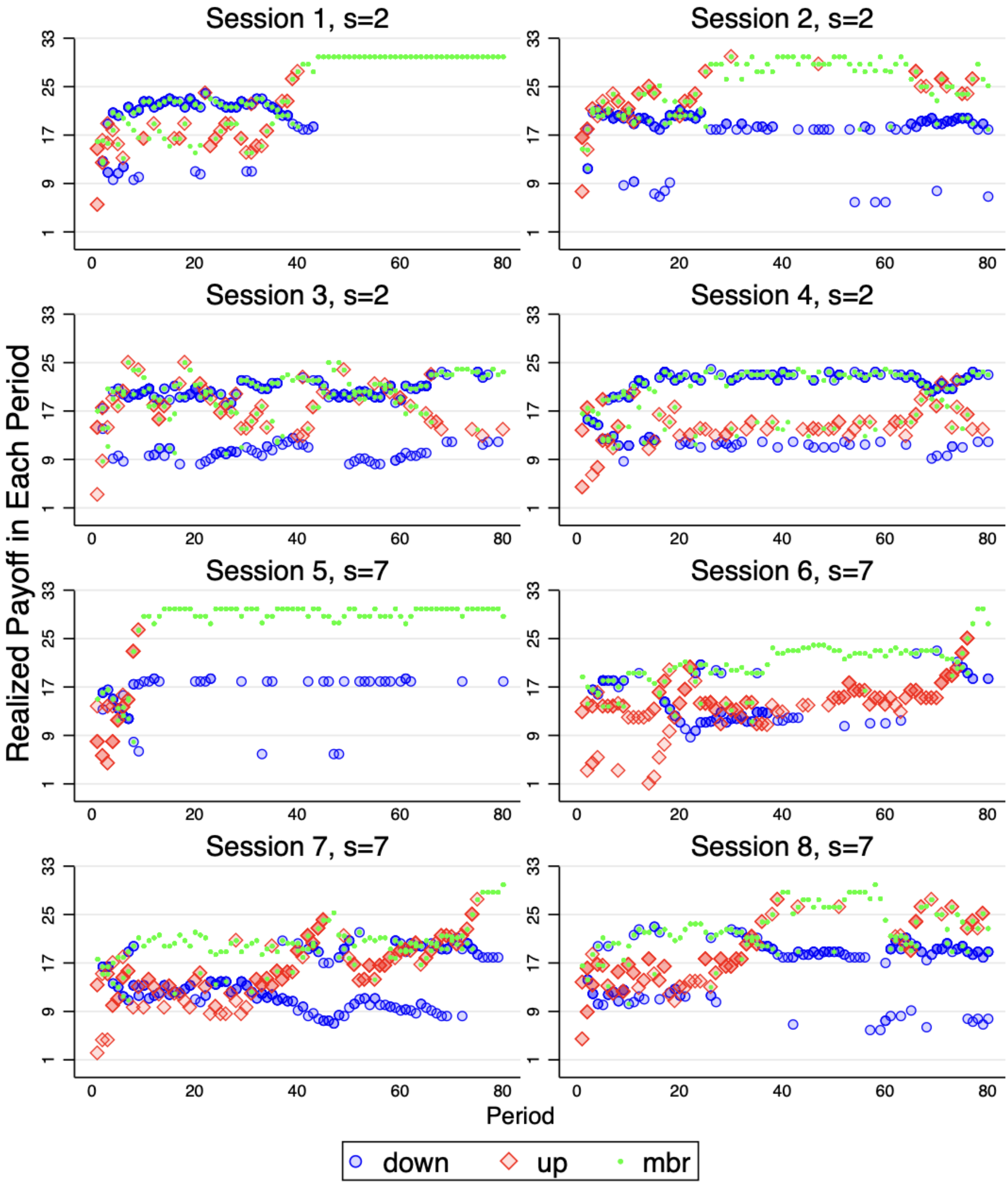


Figure. C.8: Realized payoffs of subjects playing mBR, downward and upward deviations in each period.

Table C.2: Summary statistics. Mean (SD).

Session	1	2	3	4	5	6	7	8	Total	N
Female	0.43 (0.51)	0.57 (0.51)	0.71 (0.47)	0.93 (0.27)	0.46 (0.52)	0.85 (0.38)	0.69 (0.48)	0.71 (0.47)	0.67 (0.47)	109
Study Level	3.00 (0.00)	3.07 (0.27)	3.00 (0.00)	3.93 (0.47)	3.00 (0.00)	3.79 (0.58)	3.36 (0.63)	4.00 (0.39)	3.39 (0.56)	112
Economics	0.57 (0.51)	0.50 (0.52)	0.64 (0.50)	0.07 (0.27)	0.50 (0.52)	0.00 (0.00)	0.07 (0.27)	0.14 (0.36)	0.31 (0.47)	112
Age	19.93 (1.07)	20.29 (1.73)	20.08 (1.12)	24.14 (2.96)	20.43 (1.45)	23.29 (4.10)	21.00 (4.40)	24.00 (1.47)	21.66 (3.07)	111
CRT Q1	1.00 (0.00)	1.00 (0.00)	0.93 (0.27)	1.00 (0.00)	0.93 (0.27)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	0.98 (0.13)	112
CRT Q2	0.64 (0.50)	0.64 (0.50)	0.43 (0.51)	0.50 (0.52)	0.71 (0.47)	0.71 (0.47)	0.71 (0.47)	0.43 (0.51)	0.60 (0.49)	112
CRT Q3	0.93 (0.27)	0.86 (0.36)	0.79 (0.43)	0.71 (0.47)	0.79 (0.43)	0.79 (0.43)	0.93 (0.27)	0.71 (0.47)	0.81 (0.39)	112
CRT Q4	1.00 (0.00)	1.00 (0.00)	0.93 (0.27)	0.93 (0.27)	0.93 (0.27)	1.00 (0.00)	0.86 (0.36)	0.79 (0.43)	0.93 (0.26)	112
CRT Q5	0.93 (0.27)	0.64 (0.50)	0.71 (0.47)	0.93 (0.27)	0.79 (0.43)	0.79 (0.43)	0.86 (0.36)	0.64 (0.50)	0.79 (0.41)	112
CRT Q6	1.00 (0.00)	0.79 (0.43)	0.86 (0.36)	0.79 (0.43)	0.93 (0.27)	0.64 (0.50)	0.79 (0.43)	0.79 (0.43)	0.82 (0.38)	112
CRT Q7	0.86 (0.36)	0.71 (0.47)	0.57 (0.51)	0.57 (0.51)	0.43 (0.51)	0.50 (0.52)	0.57 (0.51)	0.36 (0.50)	0.57 (0.50)	112
CRT Q1-Q7	6.36 (0.84)	5.64 (1.50)	5.21 (1.58)	5.43 (1.45)	5.50 (1.34)	5.43 (1.16)	5.71 (1.44)	4.71 (1.86)	5.50 (1.45)	112
Seconds spent on CRT	268.14 (85.32)	237.00 (102.93)	238.43 (96.85)	301.36 (98.14)	259.00 (102.63)	286.00 (104.47)	310.14 (107.65)	298.21 (118.45)	274.79 (102.70)	112
(Q1-Q7)/Seconds*100	2.74 (1.32)	2.95 (1.70)	2.78 (1.88)	2.04 (1.14)	2.61 (1.58)	2.24 (1.13)	2.13 (1.01)	1.76 (0.94)	2.40 (1.39)	112
Risk	4.93 (2.92)	3.86 (3.25)	5.07 (3.17)	3.29 (2.61)	4.57 (3.01)	4.79 (3.42)	5.07 (2.97)	4.93 (3.36)	4.56 (3.06)	112
Trust	5.29 (2.76)	5.36 (2.53)	5.57 (2.03)	5.86 (2.03)	6.00 (2.42)	6.00 (2.99)	5.29 (2.23)	6.93 (1.59)	5.79 (2.34)	112
Patience	17.10 (8.19)	14.44 (7.96)	14.95 (7.24)	12.05 (8.30)	15.44 (7.81)	14.56 (7.60)	15.18 (8.21)	13.08 (8.14)	14.60 (7.82)	112
Negative Reciprocity	4.05 (1.59)	4.44 (1.87)	4.14 (2.11)	4.92 (2.21)	5.52 (1.66)	4.34 (2.04)	4.35 (1.83)	4.46 (2.04)	4.53 (1.92)	112
Positive Reciprocity	13.33 (2.72)	13.20 (2.29)	13.93 (1.83)	15.06 (1.26)	13.35 (2.70)	13.81 (3.13)	14.92 (1.75)	14.11 (3.32)	13.96 (2.48)	112
Altruism	0.72 (0.97)	0.81 (0.78)	0.71 (0.58)	0.83 (0.54)	1.04 (0.72)	0.69 (0.46)	0.98 (0.79)	0.82 (0.83)	0.82 (0.71)	112
Level k	3.93 (2.40)	5.00 (2.72)	4.21 (3.12)	5.43 (3.03)	3.50 (3.23)	5.86 (2.63)	5.71 (2.95)	4.71 (3.36)	4.79 (2.96)	112

Notes: Female takes missing value if subjects chose to not say. Study Level = 1 if below college, = 2 if with college degree, = 3 if undergraduate, = 4 if postgraduate (taught), = 5 if postgraduate (research). One observation of Age is replaced by missing value due to unreasonably high value (i.e., 222). See Frederick (2005); Toplak et al. (2011) for CRT, Eckel and Grossman (2002); Reynaud and Couture (2012) for risk preferences, Falk et al. (2016, 2018) for social preferences, Arad and Rubinstein (2012) for level k.

Table C.3: Robustness test of treatment effects.

	(1) Up_i	(2) $Down_i$	(3) FL_i
s7	0.493	-0.522**	0.918*
	(0.492)	(0.223)	(0.501)
Study Level	-0.577	0.170	0.250
	(0.610)	(0.314)	(0.570)
Female	-0.590**	0.614***	-0.442
	(0.301)	(0.211)	(0.404)
Economics	0.580***	0.161	-0.207
	(0.167)	(0.279)	(0.546)
nCRT	0.154*	-0.108*	0.414**
	(0.086)	(0.064)	(0.165)
Age	0.083	-0.006	-0.041
	(0.096)	(0.055)	(0.144)
Risk	0.041	-0.020	0.098
	(0.047)	(0.032)	(0.099)
Trust	0.033	-0.042	0.175**
	(0.037)	(0.041)	(0.077)
Patience	-0.013*	-0.015**	0.021
	(0.007)	(0.007)	(0.028)
Negative Reciprocity	0.000	-0.032	0.138
	(0.065)	(0.038)	(0.121)
Positive Reciprocity	-0.015	-0.009	0.059
	(0.055)	(0.014)	(0.066)
Altruism	0.160	0.171**	-0.106
	(0.207)	(0.075)	(0.278)
Level k	0.062**	-0.002	0.137*
	(0.032)	(0.039)	(0.071)
Constant	-1.943	-1.511*	-5.597**
	(1.212)	(0.782)	(2.528)
N	108	107	108

Notes: Columns (1) and (2) are fractional logit regression and column (3) is logit regression. Session clustered standard errors in parentheses. One outlier is omitted for column (2) (see footnote 15). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: Average treatment effects.

	(1)			(2)		
UpDownBR	Down	BR	Up	Down	BR	Up
s7	-0.081***	0.068**	0.019	-0.074**	0.070**	0.001
	(0.030)	(0.031)	(0.026)	(0.033)	(0.029)	(0.029)
Signal2	-0.285**	0.319***	-0.102***	-0.284**	0.324***	-0.113***
	(0.115)	(0.097)	(0.033)	(0.113)	(0.098)	(0.027)
Signal3	-0.243**	0.048	0.386***	-0.241**	0.043	0.393***
	(0.114)	(0.098)	(0.059)	(0.115)	(0.098)	(0.058)
Female	0.069***	-0.008	-0.122**	0.076***	-0.029	-0.091*
	(0.026)	(0.034)	(0.059)	(0.027)	(0.036)	(0.053)
Age	-0.000	-0.005	0.012	0.000	-0.002	0.005
	(0.007)	(0.008)	(0.009)	(0.008)	(0.007)	(0.009)
Study Level	0.036	-0.002	-0.070	0.035	-0.014	-0.040
	(0.041)	(0.053)	(0.065)	(0.047)	(0.044)	(0.057)
Economics	0.001	-0.032	0.067**	0.010	-0.055**	0.096***
	(0.032)	(0.024)	(0.034)	(0.036)	(0.026)	(0.036)
nCRT	-0.017**	0.007	0.020*	-0.013	0.008	0.010
	(0.008)	(0.011)	(0.012)	(0.008)	(0.011)	(0.014)
Risk				-0.002	-0.004	0.012**
				(0.005)	(0.004)	(0.005)
Trust				-0.005	0.003	0.005
				(0.005)	(0.005)	(0.008)
Patience				-0.002*	0.003**	-0.003***
				(0.001)	(0.001)	(0.001)
Negative Reciprocity				-0.004	0.001	0.006
				(0.005)	(0.007)	(0.012)
Positive Reciprocity				-0.000	-0.003	0.008
				(0.002)	(0.003)	(0.006)
Altruism				0.019*	-0.008	-0.021
				(0.010)	(0.010)	(0.019)
Level k				0.001	-0.006	0.011**
				(0.006)	(0.005)	(0.005)
N	20759	20759	20759	20759	20759	20759

Notes: Computed based on mixed logit regressions with categorical dependent variable $UpDownBR_{it}$ taking three values for best responses, downward and upward deviations. The session clustered standard error is used. The outlier is excluded (see footnote 15). Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: Robustness test of treatment effects on the responsiveness to signals.

	(1)		(2)	
Action	1	3	1	3
s7	1.443***	0.503	1.572***	0.478
	(0.457)	(0.335)	(0.542)	(0.335)
Signal2	-3.817***	-0.821***	-3.974***	-0.814***
	(0.283)	(0.192)	(0.232)	(0.189)
s7×Signal2	-0.827	-0.481	-0.788	-0.567*
	(0.895)	(0.308)	(0.954)	(0.339)
Signal3	-2.547***	3.341***	-2.709***	3.409***
	(0.551)	(0.264)	(0.530)	(0.332)
s7×Signal3	-0.987	0.419	-0.911	0.431
	(0.875)	(0.435)	(0.886)	(0.464)
Study Level	0.441	-0.299	0.342	-0.113
	(1.259)	(0.707)	(1.022)	(0.627)
Female	-0.627	-0.892***	-0.422	-0.848***
	(0.475)	(0.302)	(0.443)	(0.278)
Economics	0.178	0.418	0.346	0.467*
	(0.310)	(0.261)	(0.265)	(0.267)
nCRT	0.002	0.160***	0.078	0.154**
	(0.160)	(0.061)	(0.153)	(0.061)
Age	-0.084	0.043	-0.045	0.007
	(0.222)	(0.117)	(0.173)	(0.107)
Risk			-0.040	0.061
			(0.065)	(0.044)
Trust			-0.183***	-0.009
			(0.053)	(0.041)
Patience			-0.026	-0.010
			(0.030)	(0.008)
Negative Reciprocity			-0.022	0.030
			(0.055)	(0.068)
Positive Reciprocity			-0.036	-0.010
			(0.042)	(0.034)
Altruism			0.131	-0.057
			(0.285)	(0.124)
Level k			-0.012	0.047
			(0.075)	(0.036)
Constant	1.278	-1.145	2.544**	-1.302
	(0.954)	(0.783)	(1.230)	(0.895)
N	25920		25920	

Notes: Mixed logit regressions. Action 2 is used as base outcome. Standard errors are clustered at session level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: Average marginal effects computed based on regressions in Table C.5.

	(1)			(2)		
Action	1	2	3	1	2	3
s7	0.036**	-0.083***	0.046	0.045**	-0.081**	0.036
	(0.015)	(0.029)	(0.028)	(0.018)	(0.032)	(0.033)
Signal2	-0.219***	0.269***	-0.050**	-0.220***	0.272***	-0.052***
	(0.017)	(0.016)	(0.021)	(0.018)	(0.017)	(0.020)
Signal2 $s = 2$	-0.172***	0.223***	-0.051**	-0.166***	0.218***	-0.052**
	(0.025)	(0.023)	(0.022)	(0.027)	(0.027)	(0.021)
Signal2 $s = 7$	-0.256***	0.310***	-0.054*	-0.268***	0.323***	-0.056*
	(0.014)	(0.022)	(0.029)	(0.015)	(0.025)	(0.031)
Signal3	-0.246***	-0.265***	0.511***	-0.248***	-0.270***	0.518***
	(0.014)	(0.021)	(0.017)	(0.011)	(0.024)	(0.019)
Signal3 $s = 2$	-0.177***	-0.316***	0.493***	-0.171***	-0.331***	0.502***
	(0.021)	(0.030)	(0.019)	(0.024)	(0.037)	(0.027)
Signal3 $s = 7$	-0.295***	-0.219***	0.514***	-0.309***	-0.211***	0.520***
	(0.030)	(0.036)	(0.025)	(0.032)	(0.039)	(0.027)
Study Level	0.030	0.016	-0.046	0.021	0.000	-0.021
	(0.056)	(0.118)	(0.068)	(0.045)	(0.097)	(0.065)
Female	-0.017	0.122***	-0.104***	-0.006	0.108***	-0.102***
	(0.027)	(0.044)	(0.039)	(0.025)	(0.037)	(0.038)
Economics	0.001	-0.052**	0.050	0.010	-0.062***	0.053
	(0.021)	(0.023)	(0.039)	(0.020)	(0.021)	(0.039)
nCRT	-0.003	-0.017*	0.020***	0.001	-0.019**	0.018**
	(0.008)	(0.010)	(0.007)	(0.008)	(0.010)	(0.007)
Age	-0.005	-0.002	0.007	-0.003	0.001	0.002
	(0.010)	(0.020)	(0.011)	(0.008)	(0.017)	(0.011)
Risk				-0.003	-0.005	0.008*
				(0.003)	(0.006)	(0.005)
Trust				-0.010***	0.007	0.002
				(0.003)	(0.005)	(0.005)
Patience				-0.001	0.002	-0.001
				(0.002)	(0.002)	(0.001)
Negative Reciprocity				-0.002	-0.002	0.004
				(0.002)	(0.009)	(0.008)
Positive Reciprocity				-0.002	0.002	-0.001
				(0.002)	(0.004)	(0.004)
Altruism				0.008	0.002	-0.010
				(0.014)	(0.019)	(0.014)
Level k				-0.002	-0.005	0.006
				(0.004)	(0.005)	(0.004)
N	25920	25920	25920	25920	25920	25920

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Instructions

Please read the following instructions carefully. To ensure that you fully understand the instructions, you will be asked to take a comprehension quiz after reading the instructions. While you are answering the quiz, please feel free to refer to the paper instructions. It will be followed by a training stage to help you get familiar with the interface. If you have any questions, please raise your hand so an experimenter would come to help you in person. Please silence your phone and any other electrical devices. Communication with other participants during the experiment is not allowed. Your cooperation is much appreciated.

The Rule

In this experiment, you will earn tokens by playing a 14-player game. At the end of the experiment, the tokens you earn will be converted into money as will be described below. The greater are your token earnings, the greater are your money earnings.

All the players in the game will choose one out of three available actions, **1**, **2**, and **3**. Every player's payoff (in tokens) will depend on his/her own action choice and the other players' action choices in a group. Each player's payoff possibilities are the same, and will be presented in a payoff table, similar to the one in Table C.7 below.

Your Actions	Fractions: x_1	x_2	x_3
1	30	0	13
2	9	22	0
3	10	6	24

Table C.7: Payoff Table Example.

The table rows are labeled by your choice (**1**, **2** and **3**) and the columns are labeled by the fraction of the other players in your group (**excluding you**) choosing each action:

- x_1 is the fraction of players choosing action **1**;
- x_2 is the fraction of players choosing action **2**;
- x_3 is the fraction of players choosing action **3**;

Each table entry represents a token payoff given the indicated choice and fraction. Suppose that all the other players in your group choose action **1**, thus we have fractions $x_1 = \frac{13}{13}$, $x_2 = 0$, and $x_3 = 0$. If you choose action **1**, you will get 30 tokens, the number in the first row and first column. If all the other players choose action **2** instead of action **1**, the three fractions are $x_1 = 0$, $x_2 = \frac{13}{13}$, $x_3 = 0$ respectively. Now by choosing action **1**, you will get 0 token (the number in the first row and second column). Similarly, if all the other players choose action **3**, you will get 13 tokens by choosing action **1** (the number in first row and third column).

Now let's suppose that among **the other 13 players** in your group, there are 4 players choosing action **1**, 4 choosing action **2** and 5 choosing action **3**. Thus the fractions are $x_1 = \frac{4}{13}$, $x_2 = \frac{4}{13}$, and $x_3 = \frac{4}{13}$. Your payoff of choosing action **1** now is the weighted average of all the entries in the first row: 14.23 tokens = 30 tokens$\times \frac{4}{13}$+0 tokens$\times \frac{4}{13}$+13 token$\times \frac{5}{13}$.

How your payoff is calculated when you choose action *1* therefore can be generalised as follows:

$$30 \text{ tokens} \times x_1 + 0 \text{ token} \times x_2 + 13 \text{ tokens} \times x_3$$

Similarly, if you choose action *2* (see the entries in the second row), your payoff will be

$$9 \text{ tokens} \times x_1 + 22 \text{ tokens} \times x_2 + 0 \text{ token} \times x_3;$$

if you choose action *3* (see the entries in the last row), your payoff will be

$$10 \text{ tokens} \times x_1 + 6 \text{ tokens} \times x_2 + 24 \text{ tokens} \times x_3.$$

Your Decisions

In the experiment, you will play a 14-player game as described above for **80 periods**. Every period will last for **10 seconds**.

You will be matched anonymously with counterparts, some or all of the participants in today's experiment. The group members will stay the same throughout the whole experiment.

Before the game starts, every player in a group will be assigned an action choice, out of three available actions. You will be informed of the action assigned to you after the game starts.

Periods

At the beginning of each period, you will be provided with one piece of information:

- We will use **random procedure** to select one player among **the other 13 players** in your group (**excluding you**) for s times, and show you their action choices from the last period. **Note**:
  - Each time, the selection is from the whole group but excluding you.
  - That is, it is possible the same player in your group might be selected more than once. As the result it is possible that all of these s randomly selected decisions might come from the same player, however the chances of this are relatively small.
  - Every player in a group will receive a piece of such information. The number of selection times is the same for everyone. However, the information may differ across players in your group, and you will only see your information but not others' information.

– You will be included in the random selection for the information to others.

Figure C.9 shows an example of the information with $n = 3$. Blue represents action **1**, green is action **2**, orange is action **3**. The horizontal axis is the period. Here the stacked column is the information to you at the beginning of period 12: among 3 times of selections, every action is observed once, namely $\frac{1}{3}$ labelled on the column. Note: the more players choosing an action, then the action is more likely to be observed in each selection.

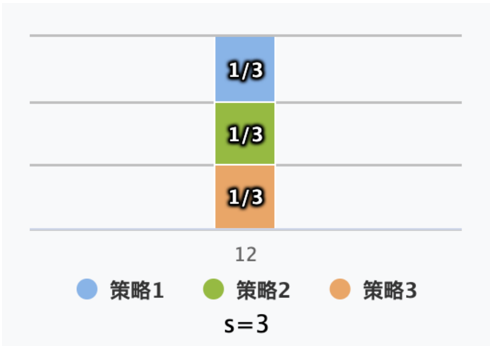


Figure. C.9: Information Example ($s = 3$).

After you review the information provided, you can choose one from three actions. You have to click 'Submit' button to successfully submit your choice. You can also decide to not make any choice. If you do not choose before a period ends, your action in the last period will be carried over. You can only make at most one choice in each period. At the end of a period, we will calculate each player's payoff of the current period using the fractions based on the players' actions at this moment.

The procedure is the same in every period. Your payoff for playing the game will be sum of payoffs across all the periods. You will not be informed of the payoffs until the end of the experiment.

Note: The information shown to you at the beginning of the first period is based on players' assigned actions. The payoff table and the selection times are same for every player in your group and stays the same in all the periods. Before game starts, we will inform everyone of the selection times s. When and which action to choose are up to you. It is in your best interest to make decisions which you believe would give you the most tokens.

Survey

You will answer a survey after finishing the game.

Your Earnings

You will earn ¥5 for showing up for the experiment.

The tokens you earn in the game will be converted into yuan. 120 tokens = ¥1.

At last, you will earn some amount of money for answering the survey. The details will be made clear as it comes.

Your total earnings of participating in this experiment will be the following:

Total Earnings = ¥5 for showing up + token(s) earned in games/120 + earnings for survey

C.4 Screenshot of Game Interface

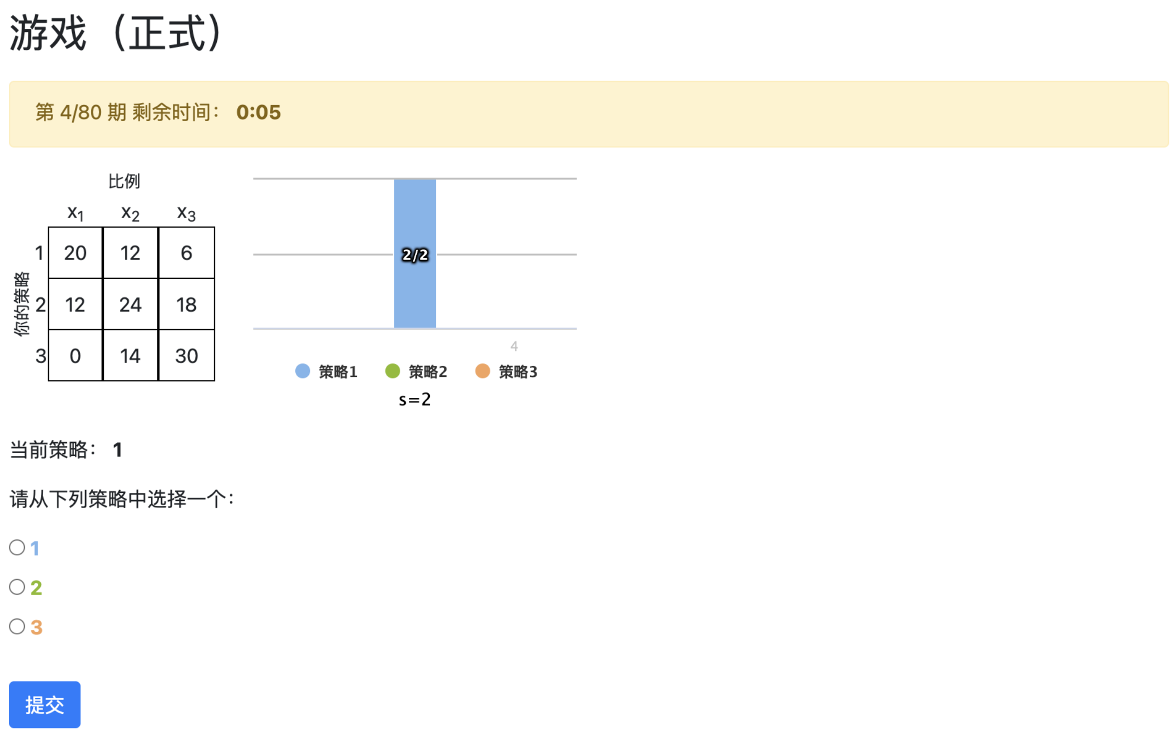


Figure. C.10: Interface.