Mixing Propensity and Strategic Decision Making*

Duk Gyoo Kim† Hee Chun Kim ‡

December 22, 2017

Abstract

This paper examines a link between an individual's strategic thinking in beauty contest games and (possibly non-rational) decision-making patterns in a non-strategic setting. Experimental evidence shows that subjects' strategic behavior, which used to be understood as a result of (possibly limited) cognitive iterations, is closely related to the non-strategic decision-making patterns. We claim that such a relationship partially explains conflicts of the previous reports on the strategic behaviors observed in the laboratory. We require attention to this relationship in that the assumption that individuals are rational in the decision-theoretic sense may create sizable misinterpretation of strategic behavior.

1 Introduction

A growing number of studies in economics and political science consider bounded rationality both in a non-strategic environment where a single player makes a decision under an uncertain state,[1] and in a strategic environment where she responds to the other agents' unknown intention and behavior. When making a decision in the non-strategic environment, individuals are often cognitively limited: They may not recognize or understand all the aspects that affect their payoffs, or lack a cognitive ability to draw an ideal decision as much

*We thank Colin Camerer and Brian Rogers for their encouragement, and thank Sangwon Park and participants of 2016 Western Economic Association International Conference for their comments.

†Department of Economics, University of Mannheim, *d.kim@uni-mannheim.de*

‡Korea Information Society Development Institute, *heechunkim@kisdi.re.kr*

[1]In the sense that a decision maker may use some strategies to deal with uncertainty, the non-strategic environment does not necessarily mean it involves no strategy at all. Rather, it may be understood as a single-player game.

as they needed. Observations from the strategic environment also seem to be away from the theoretical predictions attained under the assumption of full rationality, not only because their rationality is bounded, but also because their belief about other individuals' bounded rationality varies.

Our primary goal is to examine how individuals' non-strategic—and possibly non-rational—decision-making patterns over probabilistic events are related to their strategic ones. To analyze strategic observations, the main body of the literature has implicitly assumed that "individuals are rational in the decision-theoretic sense of choosing strategies that are best responses to consistent beliefs" (Crawford (2016)), which we hereinafter call decision-theoretic rationality. However, when subjects are asked to make repetitive decisions under uncertainty, experimental work shows that a significant amount (more than 40%) of subjects do not make decisions to maximize their expected payoff, but *match* their decision frequencies to the probability of events, which is called *probability matching* (Rubinstein (2002), Neimark and Shuford (1959)). For example, when people are asked to play ten rounds of Matching Pennies (MP) games and they are informed that for each game a coin will be tossed independently and the coin lands heads with 70% of chance, some of them choose Heads for seven out of the ten rounds, and Tails for three out of the ten, to match their choice frequencies with the probability of events. To maximize the expected payoff, they should have chosen Heads for all the rounds. We introduce a broader notion of the probability matching as there could exist other decision-making patterns: Even if they choose Heads for all the ten rounds in one set, we cannot rule out the possibility for them to choose Heads for all the rounds in seven sets out of ten repetitive sets, and choose Tails for all the rounds in the other three sets out of the ten. Or, even if they knew choosing Heads all the time can maximize the expected payoff, they may still want to consistently choose an outside option when available, with hoping that will hedge. We call this broad notion of individual tendency in repetitive decision makings as a *mixing propensity*, because such a tendency will result in mixed choices in the same environment.

We claim that without considering individuals' mixing propensity, it is challenging to map individuals' strategic behaviors to their underlying belief. The beauty contest game or its modified versions have been used to estimate individuals' cognitive levels and underlying belief about the population. One of typical beauty contest games goes as follows: "Each of those who participate in this game simultaneously submits a number between 0 and 100, and a huge prize goes to the person who submitted the closest number of two thirds of the average of all the submitted numbers." The unique Nash equilibrium strategy is for everyone to submit 0, but this obviously requires sufficiently many, if not infinite, steps of iterated dominance: If everyone randomly chooses one number between 0 and 100, the average of the submitted number would be 50, that is, the winner would be someone who submits a number close to 33. If everyone who follows the same logic submits 33, then 22 would be the winning number. If everyone picks 22, then 15 would be the winning number, and so on. That is, to reach the equilibrium strategy, one must assume that everyone could follow this logical iterations as many times as needed, which is not always true. We build upon two leading theories formalizing bounded rationality in strategic thinking: the Level-k (Lk) model (Costa-Gomes et al. (2001), Costa-Gomes and Crawford (2006)) and the cognitive hierarchy (CH) model (Camerer et al. (2004)). Both models assume that individuals use only finite ($=k$) steps of iterated dominance, and such k varies by individual. One distinctive difference is that the Lk model assumes that individuals believe others' cognition level is homogeneous at L*k-1*, while the CH model assumes that they believe it is distributed over L*0* to L*k-1*, where the population distribution is assumed to be stable. To analyze experimental observations, previous studies implicitly share an assumption that every subject is equipped with decision-theoretic rationality. In other words, the main body of the literature has assumed that no subject has any sort of mixing propensity, which may create sizable misinterpretation: An individual who has a certain type of mixing propensity may show homogeneous choice patterns even when she has a heterogeneous belief, while an individual who has another mixing propensity may make heterogeneous choice patterns that fully reflect her heterogeneous belief when the best response to the belief is a probabilistic mixture of many choices. Naturally followed questions are whether the lack of consideration of mixing propensity deteriorates the previous studies' elicitation of the structure of beliefs, and if so, how severe it is.

To address our questions, we conducted two sets of separate laboratory experiments within subjects: An ordinary decision making (ODM) experiment using a modified matching pennies game, and a strategic decision making (SDM) experiment using a modified beauty contest game. In a nutshell, from the ODM experiment we identify individuals' mixing propensity and categorize it into one of four types. Within such a mixing propensity type, observations from the SDM experiment can be analyzed in a clearer way so that we can describe better the belief distribution.

Our observations are summarized as follows: First, in the ODM experiment, about a half of the subjects were classified as a Rational Optimizer (RO), who chooses rational decisions consistently, a third of the subjects were classified as a Probability Mather (PM), who exhibits probability matching propensity, and the rest of them were as a Hedging Mather (HM), who consistently prefers a hedging option whose payoff is adjusted downward by their own risk preferences. There was no Uniform Matcher (UM) who matches the frequencies of the choice bundles to the probability of events. Second, the overall cognitive iteration level of the PM-type subject was lower than that of the RO-type subject, and the PM-type made more various strategic decisions than the RO-type did. Third, the HM-type subjects, similar to the PM-type subjects, carried over a smaller number of cognitive iterations to make a strategic decision than the RO-type ones, but their decisions were less variant than those of the PM-type subjects. Altogether these observations clearly suggest that the estimated population distribution of cognitive ability to carry over strategic thinking has been distorted *downward* in the previous studies, because of the existence of a substantial fraction of population, the PM-type and the HM-type subjects, who systematically deviate from the decision-theoretic rationality. We also claim the PM-type subjects can represent their entire underlying belief structure more accurately, and the HM-type and RO-type subjects' cognition level is likely to be underestimated.

The rest of this paper is organized as follows. In the following subsection we review related studies. Section 2 describes details of experimental design. The statistical method for inference is described in Section 3. Section 4 shows results of experiment and discusses its implication. Section 5 concludes.

1.1 Related Literature

This study is developed on the empirical and theoretical findings about bounded rationality in strategic behavior. Among many behavioral models that have been considered, we mainly focus on two leading models: Level-k model by Costa-Gomes and Crawford (2006) and Cognitive Hierarchy model by Camerer et al. (2004). Both models share two assumptions that (1) individuals are rational in the decision-theoretic sense of choosing strategies that are best responses to consistent beliefs and (2) individuals play strategies of a finite level of iterated dominance. They differ in the assumption about how subjects believe other players' strategic behavior. Level-k (Lk) model assumes that individuals have the uniform belief such that all their opponents play the same level of iterated dominant strategy. For example, $L2$ subject assumes that all their opponents play a one-time iterated dominant (or $L1$) strategy. From that assumption, Lk subjects are supposed to play a certain strategy that best responses to their uniform belief. In the Costa-Gomes and Crawford (2006), about 55% of subjects are identified as showing a certain level of play that can be interpreted as adopting the Level-k model. On the other hand, some subjects explicitly mixed two or more different strategies that represent different levels of iterated dominance. Such systematic pattern does not coincide to the assumption of uniform belief. Costa-Gomes and Crawford (2006) found source of such deviation from the learning. That is, even individuals started from the initial uniform belief, the experience leads subjects to shift to the higher level of iterated dominance while keeping the uniform belief structure. However, for some subjects, such mixing occurred irrelevant to the time horizon. These observations suggest us to think of alternative model that can explain such behavioral pattern. Cognitive Hierarchy (CH) model allows individuals to have heterogeneous belief structure. For example, L2 subject assumes their opponent plays not only L1 strategy but also L0, which is uniform random, strategy. Depends on the individual belief about the proportion of users of two different strategies, each subject may find a different best response. In Camerer et al. (2004), authors explicitly estimated the structure of belief by using observations from previous studies and their original experimental observation. However, even CH model allows the heterogeneous belief structure, they cannot fully explain the observations with mixed choices. That is, as the best response to the heterogeneous belief, choosing the interim choices consistently can be strictly better than mixing the different choices those correspond to different strategies respectively.

This study attempts to explain such a puzzle by relaxing the assumption that all individuals have the same (rational) response to the same belief. That is, we consider that individuals may show different response to the same belief and this difference in decision-theoretic rationality may lead to the apparent puzzle that mixes different strategies. Examples abound. In Rubinstein (2002), about a half of undergraduate subjects matched their frequency of choices to the probability of events for repetitive decision-making tasks. Thaler (2016) reports a similar result with MBA students at a top university. Though the detailed contexts varied, the fundamental question that the authors asked subjects to perform was the independent repetition of the MP game described above.[2] Likewise, many studies in psychology literature found significant propensity for mixing different strategies. Neimark

[2]Rubinstein (2002) performed "catch the messenger" game, where a detective's task is to determine the location of a video camera for each day, to identify unknown messengers as many as possible, while the probability to catch the messenger at each location is known. The video camera should have been installed all the time at one location where the probability is the highest, but only a small portion of students always played such a stochastic dominant action. Thaler (2016) asked MBA students to make a streak of 5 matching-pennies choices. Each choice is either Head or Tail, and a fair coin was tossed 5 times after they made choices. The payoff of matching at Tail was 1.5 times higher. They should have chosen Tail all the time, which is stochastically dominant, but the most common allocation was three Tails and two Heads, matching the ratio of the payoffs.

and Shuford (1959) and Vulkan (2000) also provide lab-experiment observations that support the existence of probability matching behavior. When we consider that similar probability matching behavior can occur at the strategic decision making process, the mixing strategies of different levels could be better reflection for the underlying belief structure.

Though probability matching has been well documented in the literature, few experimental studies explicitly considered such behavioral patterns in the optimization process in the identification of underlying belief structure in the strategic decision making environment. Georganas et al. (2015) considered whether individuals show the similar level of iterated dominance in the different form of the game. Georganas et al. (2015) conducted several different (strategic) games to the same individuals. Specifically, subjects practiced four different non-strategic tests and played a strategic decision making session. In the strategic decision making session, subjects played the 'undercutting game' and 'beauty-contest game' for four and five times respectively. While the undercutting game only allowed the discrete choices, the beauty contest allowed some interim choices that do not represent any level of iterated dominance. In the result, even in those two games shared the similar structure that requires players to exploit iterated dominant strategy, individuals show almost no correlation between the levels of iterated dominance. Moreover, there was no significant connection was found between the individual trait, such as IQ, and the level of iterated dominance. Georganas et al. (2015) attempted to find a consistency of the strategic process in different environment, but did not considered it in terms of individual optimization pattern.

2 Experimental Design

We use a within-subject design. The same subjects participated in two different experiments: In the ordinary decision making (ODM) experiment, subjects made a streak of decisions of which payoffs depend on unknown but realized events. In the strategic decision making (SDM) experiment, subjects made a streak of decisions of which payoffs depend on the randomly matched subject's decisions.

2.1 Overview

In the ODM experiment, subjects repeatedly play modified Matching Pennies games with unknown events by themselves. See Table 1 for illustration. Subject's options are on the first column; U, M and D in this example. The first row shows events and probabilities; (L=3/4, R=1/4), in this example, means that event L will be realized with probability 3/4, and event R otherwise. The matrix shows the subject's payoff. For example, if she chooses M and event L is randomly drawn, she earns $(3 - v)/4$ points, where v is for the subject's certainty equivalent. v was separately measured.

Game 1	L= 3/4	R= 1/4
U	1	0
M	$3(1-v)/4$	$(1-v)/4$
D	0	1

Table 1: An Example of the Payoff Matrix in the ODM Experiment

The ODM experiment consists of four separate games and each game consists of four sets respectively. Each set also consists of four rounds. Therefore subjects make decisions in a total of 64 rounds. A new event is drawn from the known probability distribution at the beginning of each set (four rounds). Subjects are informed that an event (either L or R in the example) is realized, and that event will not be changed within a set, but they do not know which event is realized. That is, subjects face the same but unknown event for four rounds. After that, a new event is drawn, and they face another unknown event for another four rounds, and so on. Based on subjects' choice patterns from four different games, we categorize their mixing propensities.

The SDM experiment was conducted with the same subjects who participated in the ODM experiment. In the two-player beauty contest game in Costa-Gomes and Crawford (2006), subjects earn more when they guess the match's action more accurately. This idea is maintained in our SDM experiment. Both player 1 (P1) and player 2 (P2) know the choice intervals and target parameters of P1 and P2. P1's goal is to submit a number within P1's choice interval, and P1's payoff is larger as the number is closer to P2's number times P1's target parameter. Similarly, P2's goal is to choose a number within P2's choice interval, and P2 earns more if the number is closer to the P1's number times P2's target parameter. For example, if both players are choosing a number between 0 and 100, and the winner goes to the person who submits a closer number of two thirds of the other player's number, P1's range is [0,100], P1's target parameter is 2/3, P2's range is [0,100], and P2's target parameter is 2/3. Three distinctive differences between ours and the previously conducted two-player beauty contest games are as follows: (1) Subjects play eight rounds of beauty contest games in five sets. At each set, they play with a new match. This setup allows us to fully understand the individuals' mixing propensity type. (2) The payoff function is deliberately designed in a way to distinguish a player's deterministic choice from a naïve random choice within an interval. (3) We introduce a calculation panel which tracks subjects' exact thought process.

2.2 The ODM Experiment

We design the ODM experiment to identify an individual's mixing propensity type. The entire ODM experiment consists of four different Matching Pennies games, and subjects play each game repetitively. Each game consists of four sets and each set consists of four rounds. That is, each Matching Pennies game is repeated for 16 rounds. Subjects are told that a new event is randomly drawn from the known probability distribution per each set. Subjects face the same event for four rounds within a set and as the set changes they will face another event for another four rounds. Since there are four different games, each subject plays 64 rounds (4 games $\times$ 4 sets $\times$ 4 rounds) during the entire experiment.

To prevent subjects' learning about the event from previous outcomes, the outcome of the game was not disclosed to the subjects during the experiment. They were informed of the realized outcome at the end of the experiment and got paid privately according to the outcome. Moreover, the game with the events (of computer player) allows us to prevent them from concerning the other-regarding preferences, such as inequity aversion (Fehr and Schmidt, 1999).

Table 2 describes four Matching Pennies games used in the ODM experiment. In every round, subjects choose one entity among those in the first column. The event is drawn from the first row with the probability associated with each event. For example, subjects can choose one among U, M and D in Game 1 and an event is either L with probability 3/4 or R with probability 1/4. Subject's payoff is described in the payoff matrices.

We varied the structure of each Matching Pennies game by (1) the existence of dominant actions, (2) the number of choices and (3) the highest expected payoffs subjects can earn. Table 3 summarizes the structure.

Note that M in Games 1 and 2 and B in Games 3 and 4 are choices that subjects can earn nonzero payoffs for any event, which we call a hedging action. The discount of payoffs for the hedging action, v_i for subject i is adopted to prevent the subject's bias toward the hedging action due to risk aversion. Since the hedging action gives exactly the same expected payoff from each single action choice and always guarantees a positive amount of payoff, risk-averse subjects may consider the hedging action as the dominant choice as long as the payoff is greater than the certainty equivalence of the game. To exclude this concern, we measured

Game 1	L = 3/4	R = 1/4
U	1	0
M	$3(1-v_i)/4$	$(1-v_i)/4$
D	0	1

Game 2	L = 1/2	C = 1/4	R = 1/4
U	1	0	0
M	0	1	0
D	0	0	1
B	$(1-v_i)/2$	$(1-v_i)/4$	$(1-v_i)/4$

Game 3	L =1/2	R = 1/2
U	1	0
M	$(1-v_i)/2$	$(1-v_i)/2$
D	0	1

Game 4	L = 1/4	LC = 1/4	RC = 1/4	R = 1/4
U	1	0	0	0
MU	0	1	0	0
MD	0	0	1	0
D	0	0	0	1
B	$(1-v_i)/4$	$(1-v_i)/4$	$(1-v_i)/4$	$(1-v_i)/4$

Table 2: Matching Pennies Games in the ODM

Each subject plays all four games in random order. Discount for a hedging behavior, v_i (Game 1 and 2: M, Game 3 and 4: B), varies by subject's risk preferences, which was measured by a survey.

their risk preferences before the beginning of the ODM experiment, and discounted their payoff of the hedging action accordingly so that the payoff of the hedging action equals the certainty equivalence of the game.[3]

	Existence of Dom. actions?	The number of states	The highest expected payoff
Game 1	Y	2	3/4
Game 2	N	2	1/2
Game 3	Y	3	1/2
Game 4	N	4	1/4

Table 3: Comparisons of Four Matching Pennies Games

From the ODM experiment, we categorize subjects into one of the four possible types, based on their choice patterns: The Rational Optimizer (RO) plays an optimal action that maximizes the expected payoff for all rounds for all sets. The Probability Matcher (PM) mixes his/her action to match the given probability within each set and this mixing proportion is equal across the sets. The Uniform Matcher (UM) plays the same action within each set but the proportion of sets with a single action will be equal to the given probability. The Hedging Matcher (HM) plays an 'intermediate' action for all rounds at all sets. We use Maximum Likelihood estimation for categorization. Those four types are distinguished by their behavioral pattern.

Theoretical Benchmark 1. *(Mixing Propensity) Individuals with a different mixing propensity are to show different decision-making patterns;*

1. *A RO-type player always chooses the action that maximizes the expected payoff.*
2. *A PM-type player mixes different actions within each set and the proportion of mixing follows the probability distribution of the corresponding states.*
3. *A UM-type player chooses a single action within each set, but changes the action across the sets. The proportion of such mixing follows the probability distribution of the corresponding states.*

[3]We followed Holt and Laury (2002) to measure the risk preferences.

4. A HM-type player always chooses an hedging action that provides a positive payoff in any cases.

Each type has a different choice pattern in the set and the game-wise level. The following table 4 shows possible choice patterns of each type in Game 1.

Game 1	Set 1	Set 2	Set 3	Set 4
RO	U4	U4	U4	U4
PM	U3D1	U3D1	U3D1	U3D1
UM	U4	D4	U4	U4
HM	M4	M4	M4	M4

Table 4: Predicted Behavior of Four Types in Game 1

This table shows how a player with a certain type of mixing propensity will choose actions. When a player is expected to choose an action A$\in$ {U,M,D} for n times, it is denoted by An. The RO-type subject will play U, the choice that gives the largest expected payoff, all the time. The PM-type subject will match the frequency of her choices with the probability of events, so in each set of four rounds, she will mix three Us and one D, up to permutations. The UM-type subject will play the same action within a set, but she will match the frequency of her choice blocks with the probability of events. Three sets of Us and one set of Ds will be chosen, up to permutations. The HM-type subject will choose H all the time.

The RO-type subject is expected to play action U all the time since U maximizes the expected payoff. The PM-type subject is expected to play action U three times in each set because the PM-type is expected to mix his/her play to match with the given probability within each set. The UM-type subject is expected to play the same action, either U or D, within each set but the proportion of sets that each action is played will be matched to the given probability. The HM-type subject is expected to play the intermediate action M all the time.

2.3 The SDM Experiment

We design the SDM experiment to identify individual strategic decision-making patterns. The entire experiment consists of eight sets and each set consists of five rounds of the two-player beauty contest game. In each set, two subjects are randomly and anonymously matched, and play a game for a whole set of five rounds with the partner. That is, they repeat playing one beauty contest game for five rounds within a set. As the set changes, each subject is randomly re-matched to another partner anonymously, and play a new beauty contest game. Eight games have different structures in terms of players' choice intervals and target parameters.

Similar to the ODM experiment, the realized outcome from their choices was not disclosed during the experiment. That is, subjects played the game without feedback and the final outcome of their choices was informed at the end of entire SDM experiment. This restriction prevented subjects from learning in a retrospective or an experience-based way. Subjects earned payoffs at each round according to a "payoff function"[4] and the monetary compensation was paid according to sum of payoffs at the end of experiment. Every detail of the SDM experiment was informed to the subjects.

We used eight different two-player beauty contest games of which choice intervals and target parameters vary. The form of each game is described by four factors; a choice interval and a target parameter of player 1, and those of player 2. For notational simplicity, α denotes a choice interval $[100, 500]$, β denotes $[100, 900]$, δ denotes $[300, 900]$, and γ denotes $[300, 500]$. A target parameter is denoted as $n_1 = 0.5$, $n_2 = 0.7$, $n_3 = 1.1$, and $n_4 = 1.5$ respectively. Moreover, we designed the games to vary in terms of the number of iteration to arrive to Nash equilibrium choice, the pattern of iterated strategies, and the location of Nash equilibrium choice. Thus, always choosing the largest or the smallest number in the choice interval does not maximize their payoffs. This fact was also notified to the subjects at the instruction stage. Combined with the no-feedback policy, such variations prevented subjects from routinizing their strategic/behavioral choice patterns. This way, we can expect subjects to concentrate on their own strategy and belief to maximize their payoff at each game. Table 5 summarizes the details of the structure of games.

We designed eight games to be paired into four pairs so that each subject can be assigned to play the two different positions of the same game. For example, the game labeled $\alpha n_2 \beta n_4$ and the game labeled $\beta n_4 \alpha n_2$ are paired, so player 1 in game $\alpha n_2 \beta n_4$ plays the exactly same role of player 2 in game $\beta n_4 \alpha n_2$. Similarly, player 1 in game $\beta n_4 \alpha n_2$ plays the exactly same role with player 2 in game $\alpha n_2 \beta n_4$. This feature is also informed to the subjects. To minimize experience-based learning, we shuffled the order of the eight games without allowing two paired games to be played consecutively.

The eight games differ in various aspects. The main difference comes with the target structure, which describes the pair of target numbers. In two games, both target numbers are greater than 1. In another two games both numbers are smaller than 1. The other four

[4] To minimize a concern about miscalculation and heterogeneity in comprehensibility, we did not provide the exact functional form to subjects. Section 2.3.1 discusses it in detail.

Game	Target Structure	#Iterations	Pattern of Iterations	End with Dominance
$\alpha n_2 \beta n_4$	Mix/High	17	A	Y
$\beta n_4 \alpha n_2$	Mix/High	18	A	N
$\delta n_3 \beta n_1$	Mix/Low	4	A	Y
$\beta n_1 \delta n_3$	Mix/Low	5	A	N
$\beta n_1 \beta n_2$	Low	4	S	Y
$\beta n_2 \beta n_1$	Low	4	S	Y
$\delta n_3 \gamma n_3$	High	2	A	N
$\gamma n_3 \delta n_3$	High	2	A	Y

Table 5: Forms of the Beauty Contest Games

The form of each game is described by four factors; a choice interval and a target number of player 1, and those of player 2. α denotes a choice interval $[100, 500]$, β denotes $[100, 900]$, δ denotes $[300, 900]$, and γ denotes $[300, 500]$. A target number is denoted as $n_1 = 0.5$, $n_2 = 0.7$, $n_3 = 1.1$, and $n_4 = 1.5$ respectively. Target Structure describes the pair of target numbers. 'High' (respectively, 'Low') describes the case where both target numbers (for example, n_2 and n_4 in $\alpha n_2 \beta n_4$) are greater (resp., smaller) than 1. 'Mix/High' (respectively, 'Mix/Low') describes the case where one number is greater than 1 and the other one is smaller than 1, and the multiplication of two numbers is greater (resp. smaller) than 1. #Iterations describes how many steps of the iterated dominance is required to arrive to Nash equilibrium. Pattern of Iterations describes whether the number correspond to each step of the iterated dominance is monotone increasing/decreasing (denoted as 'S') or alternatively changes (denoted as 'A'). When Nash equilibrium strategy is to choose the boundary of choice interval, such strategy coincides with the iterated dominance.

games go with one target number greater than 1 and the other one smaller than 1. In two games (resp., the other two games) among those four, the multiplication of two numbers is greater (resp., smaller) than 1. The games also differ in terms of the required number of steps using the iterated dominance to arrive to the Nash equilibrium choice. In two games, the best response is monotone increasing/decreasing in the level of the iteration of dominance, while in the other six games the best responses oscillate. In five games, Nash equilibrium choices coincides the boundary of the choice interval, while Nash equilibrium choices are interior for the other three games.

2.3.1 Calculation Panel

The possibility of miscalculation is one of the main concerns about interpreting observations from the two-player beauty contest games. Previous studies consider a symmetric payoff function of which payoff decreases in the absolute difference between the actual choice and the ideal guess (the matched player's number times the target number). Though the payoff structure is claimed to be simple, it could be problematic if subjects' calculation ability varies. To avoid potential issues on miscalculation, we provide a calculation panel instead of asking subjects calculate their optimal choices based on their beliefs.

Subjects can use the calculation panel to find the exact number that corresponds to their best response for partner's choice. There are two modules in at each game's calculation panel. The module ("module A") in the upper half gives the best response for the player himself/herself with respect to the prediction of the opponent's choice. The module ("module B") in the lower half provides the best response for the opponent with respect to the prediction of the opponent's prediction about the player's choice. For each module, subjects can choose an interval, with a singleton allowed. In module A, subjects can input the lower bound and the upper bound of an interval within which their opponent's choice might be placed. After that, clicking 'Calculate' button generates the red-colored number that maximizes his/her own payoffs. The distribution over the range is fixed as the uniform distribution over the selected range and subjects were informed of it. Figure 1 considers player 1 at the game $\delta n_3 \beta n_1$. If player 1 wants to find L1 strategy with belief of partner's choice would be any number from 100 to 900, he/she may enter 100 for the lower bound and 900 for the upper bound and click the Calculate button. Module A will generate the result of 678 as the closest integer approximate of the best response 678.3. Subjects are informed that their usage of the panel and the result of usage will not be related to their monetary outcome at the end of session.

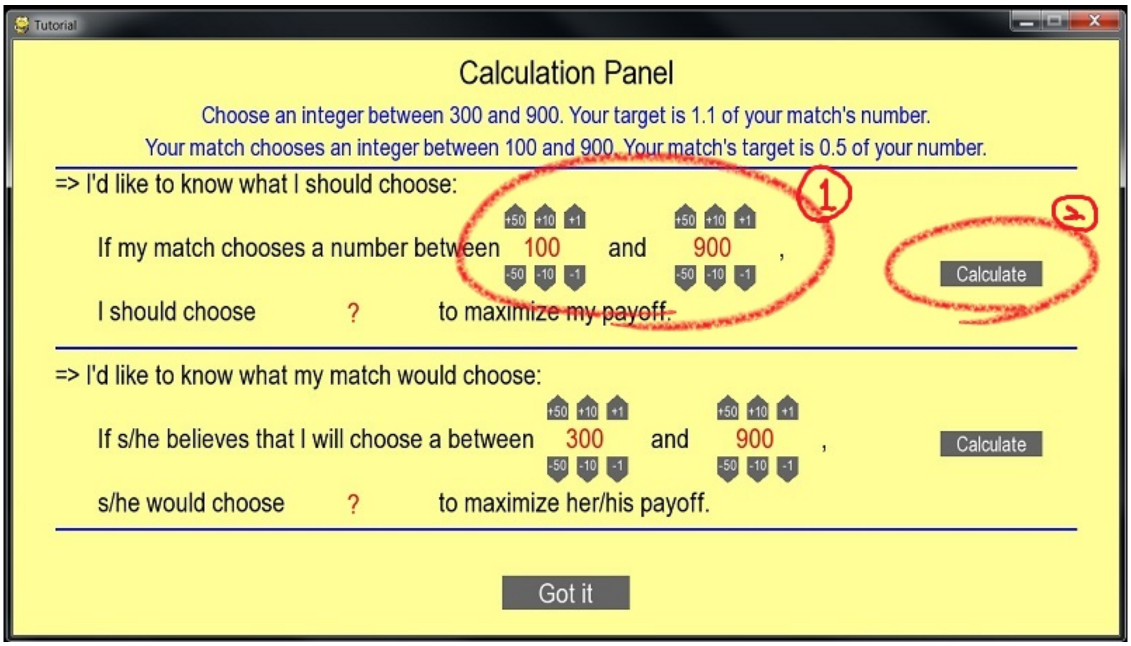


Figure 1: A Screen-shot of Calculation Panel

Two main purposes of introducing this calculation panel are to avoid potential miscalculations, and to obtain a richer data for understanding how a subject's thought process worked. In the module above, a subject can calculate his/her optimal choice based on his/her belief about the opponent's choice. In the module below, a subject knows what the opponent will choose within his/her own belief.

There are at least two noticeable merits. This calculation panel allows subjects to find the exact number corresponding to the higher order of iterated dominance (Lk with $k > 1$) without leaving possibility of miscalculation. To know the L3 choice, player 1 needs to know his/her partner's L2 choice in advance, which would be the best response to player 1's L1 choice. Thus, player 1 can input the lower bound 100 and the upper bound 900 which corresponds to his/her own range of choice to have his/her own L1 choice. After that, input the L1 choice to the range of module B (by inputing the same number in both lower and upper bounds) to calculate the L2 choice, the best response of the opponent. For example, 678 in game $\delta n_3 \beta n_1$ is the choice that player 1 will make if player 1 is of the L1 type. From this result, player 1 can input 678 in module B to have the result of the L2 choice of the opponent, which is 339. Then, player 1's best response can be obtained by inputting 339 in module A.

Second, the record of the calculation panel usage allows us to track the individual subject's decision process. That is, having tractable records for calculation process at each step of iterated dominance helps us to figure out the paths that subjects followed to reach the final decision. Even if their actual decision is different to the last calculation result, we are still able to understand how the subjects use their calculation process when making the strategic decision.

Given this structure, Table 6 shows player 1's choices correspond to each level of the iterated dominance, Nash equilibrium, and the remaining intervals correspond to each round of the iterated deletion, respectively. L1 is the first level of the iterated dominance when the subject considers his/her opponent to choose numbers with uniform probability over the all interval (one type of L0 players considered in Costa-Gomes and Crawford (2006) and Camerer et al. (2004)). $L2$ and $L3$ respectively are the second and third level of iterated dominance at which the subject considers their opponent to play $L1$ and $L2$ strategy. NE is Nash equilibrium choice of the game. For illustration, consider game $\alpha n_2 \beta n_4$. L1 strategy, the optimal strategy when the opponent's choice is random, is to choose 419. Then L2 strategy, the optimal strategy when the opponent chooses his/her L1 strategy, is to choose 361. Similarly, L3 strategy is to choose 440. The remaining interval that is strictly undominated in the first round of the iterated dominance is [100,450].

The introduction of the calculation panel still cannot clearly tell whether subjects make random choices. If subjects choose any number without using the calculation panel for a set, we can infer that they make random choices, but no subjects actually made such a pure random choice. It is also possible for subjects to make any choice even after using the

Game	L1	L2	L3	NE	1st Round	2nd Round	3rd Round	4th Round
$\alpha n_2\beta n_4$	419	360	440	500	100, 450	105, 500	105, 472.5	110.25, 500
$\beta n_4\alpha n_2$	515	629	540	750	150, 750	150, 675	157.5, 750	157.5, 708.75
$\delta n_3\beta n_1$	678	363	373	300	300, 900	300, 495	300, 495	300, 300
$\beta n_1\delta n_3$	330	339	181	150	150, 450	150, 450	150, 247.5	150, 247.5
$\beta n_1\beta n_2$	303	209	106	100	100, 450	100, 315	100, 157.5	100, 110.25
$\beta n_2\beta n_1$	419	212	146	100	100, 630	100, 315	100, 220.5	100, 110.25
$\delta n_3\gamma n_3$	463	550	550	550	330, 550	363, 550	399.3, 550	439.3, 550
$\gamma n_3\delta n_3$	500	500	500	500	330, 500	363, 500	393.5, 500	439.3, 500

Table 6: Player 1's Strategic Choices with respect to the Iterated Dominance
This table shows the theoretical predictions on various contingencies. The first column headed 'L1' shows the best response of player 1 when the opponent chooses any random number in the interval. The second to the fourth columns headed 'L2', 'L3', and 'NE', respectively, shows the corresponding best responses of player 1 when the opponent's cognitive iteration level varies accordingly. The fifth to the last columns show the strictly undominated intervals after the corresponding rounds of the iterated dominance.

calculation panel, but we cannot tell whether that choice is randomly made or derived after their own thought processes with taking into account the results of the calculation panel.

3 Results

Six sessions of the laboratory experiments were conducted at Missouri Social Science Experimental Lab (MISSEL) of Washington University in St. Louis with a total of 86 participants. From all participants, we collected 62 effective subjects from those who passed screening tests of both experiments. We estimated their type by using the data from the ODM experiment. Using MLE method[5], we found that RO-, PM-, and HM-type subjects share 45.2%, 25.8% and 29%, respectively, from all effective samples (Table 7). Since no single subject exhibits choice patterns consistent with UM type, we restrict our attention to RO-, PM-, and HM-types.

We relate this categorization of subjects to their behavioral pattern in the SDM experiment. Our main hypothesis is that subjects' behavioral patterns in the ODM experiment are inherited to the SDM experiment. We found the following two observations consistent with the predictions under our main hypothesis.

[5]For detailed statistical analysis, see Appendix. Roughly speaking, we calculate each subject's "error", provided that the subject has a specific type, and find the type that yields the smallest error.

Type	Count	%
RO	28	45.2
PM	16	25.8
HM	18	29.0
UM	0	0
Sum	62	100

Table 7: Overall Distribution of Mixing Propensity in the ODM experiment
Based on the choice patterns in the ODM experiment, we categorize subjects into one of four possible types.

1. RO-type subjects tend to show a higher level of cognitive iterations with a smaller variance of the choice distributions than PM-type subjects.

2. HM-type subjects are less likely to diversify their choices than PM-type subjects, and have a similar variance of the choice distributions to that of RO-type subjects.

In short, RO-type subjects seem to be more reflective than HM- or PM-type subjects, and PM-type subjects change their decisions according to their own belief distribution of the match's cognitive ability. That is, for example, if a PM-type subject believes that he plays the game with L0 player with 70% of chance and L1 player with 30% of chance, then he behaves as if he is of L1 player for the 70% of the whole games, and of L2 player for the remaining 30% of the games. That creates more variation of the choices. Meanwhile, HM-type subjects exhibit the lowest cognitive iterations among three groups, and they seem to consistently make 'belief-weighted' choices given their own beliefs.

Result 1. RO-type subjects tend to show a higher level of cognitive iterations with a smaller variance of the choice distributions than PM-type subjects.

To have a simple outlook about how many cognitive iterations each subject executed, we aggregated the subjects' choice data from the SDM experiment by their ODM type. Roughly speaking,[6] we did the followings. First, individual's choices are coded as one integer between 1 and 4, where number k corresponds to k-th cognitive iterations. For simplicity we call such codes as 'cognition levels.' Note that no decisions were coded as L0, since every subject did use the calculation panel at least once for each set. If the choices were made more than four cognitive iterations, including Nash Equilibrium choices, we coded them as 4. This

[6] For detailed procedure of obtaining the aggregate observations, please see Appendix.

way, we could find each subject's "distribution" of revealed cognition levels. Collecting such individual distributions by the ODM type, we could find the distribution of variances of cognition levels for each group.

	$E[\mu_i]$	$E[\sigma_i^2]$
RO	3.17	0.61
PM	2.71	1.03
HM	2.32	0.65

Table 8: Mean of Individual Means and Variances in Revealed Cognition Level, by Type
This table shows how the choice patterns in the SDM experiment differ by the type categorized in the ODM experiment. $E[\mu_i]$ refers to the mean of individual means of revealed cognitive iterations for making their decisions. $E[\sigma_i^2]$ refers to the mean of individual variances of the revealed cognitive iterations.

Table 8 shows two summary statistics about the individual distributions of cognitive iterations for each type. On average, RO-type subjects showed the cognition level of 3.71, and the mean of individual variances of the cognition level was 0.61. PM-type subjects showed the average cognition level of 2.71 with the average variance of 1.03. HM-type subjects showed the average cognition level of 2.32 with the average variance of 0.65. RO-type and HM-type subjects showed relatively lower variance than PM type subjects. On the other hand, RO-type subjects showed relatively higher average cognition level than other type subjects. This result briefly implies that those types of subjects categorized by the ODM experiment are pertinent to describe their behavioral patterns in the SDM experiment.

To consider such difference in a more detailed way, we conducted test of distributions between each pair of groups.

	Mean Tests			Variance Tests		
Types	RO&PM	RO&HM	PM&HM	RO&PM	RO&HM	PM&HM
P-Values (one-sided)	0.012	0.004	0.11	0.044	0.439	0.024

Table 9: Test Results and p-values Between Distributions

Table 9 summarizes how decision patterns are statistically different in two groups. We tested null hypothesis that two distributions are from the same population in their mean and variance, respectively.[7] The test results show that RO-type subjects, on average, have

[7]We used the Fisher method (F-test) to measure a statistical significance between two distributions.

a higher cognition level than both PM-type and HM-type subjects, and these differences are statistically significant at the 5% level and the 1% level, respectively. PM-type subjects have a weakly higher cognition level than HM-type subjects, but this difference is statistically insignificant. On average, PM-type subjects showed a larger variation of cognition level in the SDM experiment than both RO-type subjects and HM-type subjects, and these differences are statistically significant at the 5% level. There is no statistically significant difference in the variance of the cognition level between RO-type subjects and HM-type subjects. Altogether, we found that RO-type subjects are, on average, more likely to show a higher cognition level with a less variation than PM-type subjects.

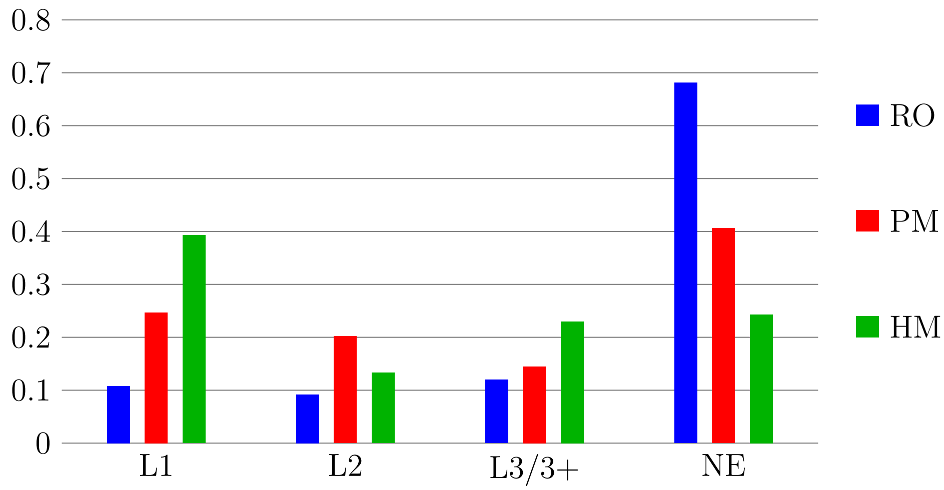


Figure 2: Proportions of Total Choices in Each Type

Figure 2 shows a proportion of choices from subjects in each type. More than 65% of choices of RO-type subjects are with a higher (4 or more) cognition level. The cognition levels of PM-type subjects are more dispersed than that of RO-type. In summary, this observation supports our presumption that PM- and RO-types inherit their behavioral pattern showed in the ODM experiment to the SDM experiment. RO-type subjects, who consistently chose actions that maximize the expected payoff in the ODM experiment, also showed consistent choice behavior at a certain level of cognition. On the other hand, PM types, who matched the frequency of their actions to a given probability of events in the ODM experiment, also spread their actions to several different cognition levels based on their belief about the cognitive level distributions of the opponent.

From Figure 2, we also find that HM-type subjects' cognition levels are as dispersed as PM-type subjects. Our second observation, in a nutshell, is that the source of such a variation is different.

Result 2. HM-type subjects are less likely to diversify their choices than PM-type subjects, and have a similar variance of the choice distributions to that of RO-type subjects.

Our second observation is that HM-type subjects are distinguished by other types in the sense that they choose a certain 'intermediate' value that may be reflected their belief distribution of the match's level of cognition. In the ODM experiment, HM-type subjects chose actions that give positive, but depreciated, payoffs in any events. We interpreted their behavior as a subjective optimization which is supposed to minimize a risk of wrong prediction. So we presumed that HM-type subjects may choose some belief-weighted value that could be calculated by a weighted average of the optimal choice on each cognition level. For example, suppose that a subject form a belief about her opponent's cognition level distribution as L0: L1: L2 = 20: 30: 50 (%). In this case, though L3 choice may give her a maximized expected payoff in any event, she may choose $0.2 * L1 + 0.3 * L2 + 0.5 * L3$ choice consistently.

To provide a statistical evidence, we test whether choice patterns of HM-type subjects are different to those of other-type subjects. Since the aggregate choice distribution of RO-type subjects is single-picked while the other two distributions are not, we focus here on the variances of individual subject' choices. In Table 8, the average of the individual variances of RO-type subjects is 0.61, and that of HM-type subjects is 0.65. We tested the null hypothesis that two sample distributions are from the population with the same variance. The test result (p-value of 0.439) cannot reject the null hypothesis. We also tested a hypothesis that these distributions have the same variance of the PM-type subject. The test result rejects the hypothesis that RO and HM type group has the same mean of the variance of individual choices of the PM type at the 5% of significance level (Table 9).

In words, both RO-type subjects and HM-type subjects, regardless of whether the choice by itself is an optimal one or not, have a tendency not to change their decisions within the same set of the SDM experiment. HM-type subjects are likely to choose an intermediate (belief-weighted) level of cognition consistently, which makes a similar variance to RO-type subjects. Their behavioral pattern is strongly distinguished from that of PM-type subjects who tend to change the choices within the same set of the SDM experiment. The observed cognition level of the HM-type subjects is lower than that of the RO-type subjects. Such a behavioral pattern is consistent with our prediction for the HM type.

3.1 Recovery of Belief Structure

We investigate the belief structure of the PM type: From the results above, we found that behavioral patterns observed in the ODM experiment have some predictive power for the SDM experiment. For RO-type subjects, this result implies that the choice distribution of RO-type subjects, which is mostly bunched at the Nash equilibrium action, cannot fully reveal their underlying belief structure. Only one thing that we can infer from this observation is that the most of RO-type subjects believe that the match will be most likely to choose the Nash Equilibrium action. However, PM-type subjects tend to diversify their responses in the SDM experiment, and their actual responses may reveal their underlying belief structure. Thus, we consider 16 PM-type subjects' actual responses to recover their underlying belief structure.

Level	L1	L2	L3/3+	NE	Unclassified	SUM
Count	97	79	57	159	248	640
w/ Unclassified (%)	15.2	12.3	8.9	24.8	38.8	100
w/o Unclassified (%)	24.7	20.2	14.5	40.6	-	100

Table 10: Overall Distribution of Cognition Level for the PM Type in the SDM

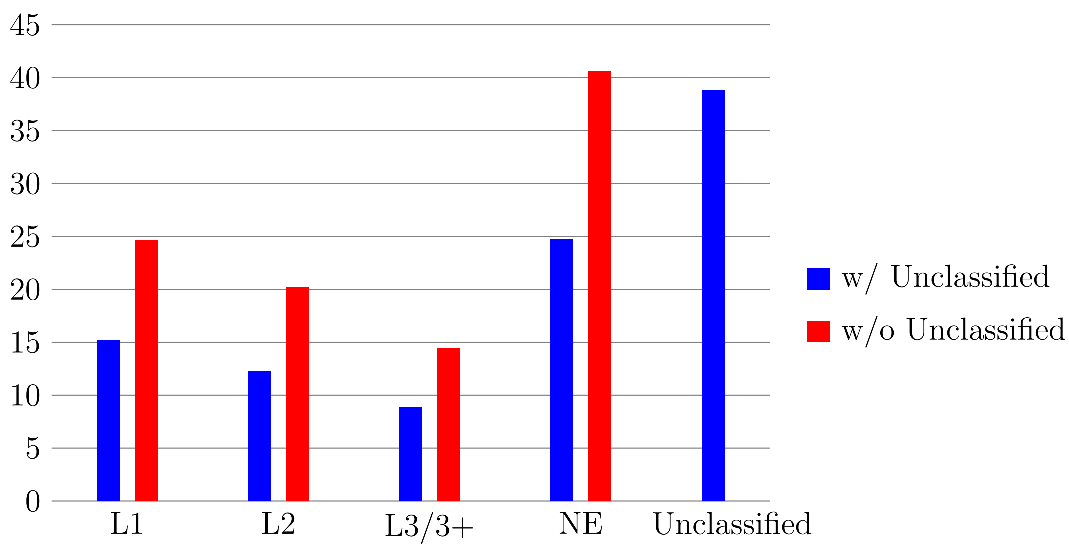


Figure 3: Proportions of Total Choices from the PM Type

Table 10 and Figure 3 summarize how PM-type subjects distributed their responses over almost every level of cognition. Even though the NE responses are the most frequently chosen actions, they allocated nontrivial proportions of their choices on the different level of cognition. The sum of proportions of L1 and L2 actions is larger than that of NE actions. This result implies that most of PM type subjects respond to the existence of L0 or L1 subjects.

Another interesting observation is that as the cognition level gets higher, the proportion of such a level gets smaller. From Table 10, proportions of L1 to L3/3+ levels are 24.7%, 20.2%, and 14.5%, respectively. When we consider unclassified responses as a random-like behavior, which correspond to L0 behavior, this trend is still consistent.

The last observation is that the proportion jumps to 40.6% at the NE action. This might be interpreted in two ways. First, PM-type subjects, compared to RO-type subjects, may believe that a smaller proportion of the population is equipped with a higher cognition level. For this reason, they may strategically assigned a smaller proportion to the higher level of cognition. Second, subjects may have some 'framing effect' on the Nash Equilibrium. That is, even though most of NE actions require higher level of cognition than other actions, they may consider it as some focal point and put the highest weight. This interpretation is also consistent with our observation for RO-type subjects: Assuming that RO-type subjects also have the similar belief structure to PM types, they are mostly best response on NE action based on their rational optimization.

4 Concluding remarks

In this study we examined how individual's (possibly non-rational) choice patterns are related to their strategic decision-making patterns. We considered each individual has a different way of making decisions when they face a probabilistic event, and categorized them into three different types: Rational Optimizer (RO), Probability Matcher (PM), and Hedging Matcher (HM). We found more than a half of entire subjects showed different choice patterns from rational optimization. Our main observation is that each type showed different decision-making patterns when they were asked to make strategic decisions. While RO-type subjects focused more on the Nash equilibrium action, PM- and HM-type subjects chose their actions responding to lower cognition levels. Especially, PM-type subjects diversified their actions to multiple different levels of cognition as they diversified their decisions in the ODM experiment.

Assuming that PM-type subjects match the frequency of the strategic choices to their belief distribution of the other players' cognition levels, as they similarly did in the ODM experiment, we can recover their underlying belief structure in a more detailed way. We observed that PM-type subjects played different actions within a set and assigned less proportion of actions to higher levels of cognition (L1: 15.2%, L2: 12.3%, L3/3+: 8.9%). This result suggests that subjects strategically assigned their actions based upon their underlying belief. Moreover, we observed PM-type subjects also assigned higher proportion (24.8%) on the Nash equilibrium action. This result supports previous result from Camerer et al. (2004) claimed that subjects consider Nash equilibrium as one of focal points. Summing up these observations, we can conclude that subjects actually have heterogeneous belief structure, which consists of several different levels of cognition, but still consider the Nash equilibrium as a plausible focal point.

The relationship between the decision-making patterns in the ODM experiment and in the SDM experiment suggest that we may have underestimated the belief of bounded rationality. If every subject were the PM type, and the belief structure estimated by the Level-k theory must be downward biased. If every subject were either the RO type or the HM type, the belief structure estimated by the Cognitive Hierarchy model would underestimate the variance of the distribution.

References

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong, "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

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.

Costa-Gomes, Miguel, Vincent P Crawford, and Bruno Broseta, "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 2001, *69* (5), 1193–1235.

Crawford, Vincent P., "New Directions for Modelling Strategic Behavior: Game-Theoretic Models of Communication, Coordination, and Cooperation in Economic Relationships," *Journal of Economic Perspectives*, November 2016, *30* (4), 131–150.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 1999, pp. 817–868.

Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber, "On the persistence of strategic sophistication," *Journal of Economic Theory*, 2015, *159* (Part A), 369–400.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, December 2002, *92* (5), 1644–1655.

Neimark, E D and E H Shuford, "Comparison of predictions and estimates in a probability learning situation.," *Journal of Experimental Psychology*, 1959, *Vol 57(5)*, 294–298.

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

Thaler, Richard H., "Behavioral Economics: Past, Present, and Future," *American Economic Review*, July 2016, *106* (7), 1577–1600.

Vulkan, Nir, "An Economist's Perspective on Probability Matching," *Journal of Economic Surveys*, 2000, *14* (1), 101–118.

Appendix A. Statistical Model Specification

A1.1. ODM Model Specification

In the ODM experiment, subjects make one choice at each round, 64 choices in total. To categorize the individual subject's decision-making patterns, we use the maximum likelihood method.

Let $x_{g,j}^{i,k}$ denote the number of subject i's decision that is identical to the type k subject's decision in game g of set j. We define $k \in K = \{RO, PM, UM, HM\}$ and $g, j = 1, 2, 3, 4$. We similarly define a vector $x_g^{i\,k} = (x_{g,1}^{i,k}, \cdots, x_{g,4}^{i,k})$.

We define $\epsilon^k \in [0, 1]$ the type specific rate of random choice that is independently and identically distributed. c_g is the number of actions that each subject has in game g. that is, $c_1 = c_2 = 3$, $c_3 = 4$, and $c_4 = 5$. Since ϵ^k is assumed to be type-specific and identically distributed over all the choices, we formulate the probability that the subject of the type k

makes some predicted decisions at game g as $1-\epsilon^k+\epsilon^k/c_g = 1-(c_g-1)\cdot\epsilon^k/c_g$.[8] Then, $L_g^{i,k}(\epsilon^k|\, x_g^{i,k})$ is the probability of observing $x_g^{i,k}$ when the subject i is of type k:

$$L_g^{i,k}(\epsilon^k|x_g^{i,k}) = \prod_{j=1}^{4}\left[1-(c_g-1)\cdot\epsilon^k/c_g\right]^{x_{g,j}^{i,k}} \times \left[\epsilon^k/c_g\right]^{4-x_{g,j}^{i,k}}.$$

Similarly, we define $\hat{x}_g^{i,k}$ the number of sets of subject i's decision that equals to the type k's decision in game g. That is, $\hat{x}_g^{i,k}$ counts the number of sets in each vector $x_g^{i,k}$ such that each set has exactly the same number of the type k subject's decision. For example, consider a RO-type subject who may choose the same action in 3 sets (say, set 1, 2, and 3) and mixed two actions in another set (set 4). Then, $\hat{x}_g^{i,k}=3$ since the number of sets that equals to the RO-type subject's decision is 3 (set 1, 2, and 3). With the similar notion, we define a vector $\hat{x}^{i,k}=(\hat{x}_1^{i,k},\,\cdots,\,\hat{x}_4^{i,k})$. Then, we define $\hat{L}^{i,k}(\epsilon^k|\,\hat{x}^{i,k})$ the probability of observing $\hat{x}^{i,k}$ when the subject i is of type k :

$$\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k}) = \prod_{g=1}^{4}\left[L_g^{i,k}(\epsilon^k|x_g^{i,k})\right]^{\hat{x}_g^{i,k}} \times \left[1-L_g^{i,k}(\epsilon^k|x_g^{i,k})\right]^{4-\hat{x}_g^{i,k}}.$$

Next, we define $z^{i,k}$ as a type indicator for subject i, where $z^{i,k}=1$ if subject i is of type k and $\sum_{k\in K} z^{i,k}=1$. From $\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k})$, subject i's maximum likelihood function can be calculated :

$$\begin{aligned}
L^i(\epsilon, z^i|x^i) &= \prod_{k\in K}\hat{L}^{i,k}(\epsilon^k|\hat{x}^{i,k})^{z^{i,k}} \\
&= \prod_{k\in K}\left[\prod_{g=1}^{4}\left\{L_g^{i,k}(\epsilon^k|x_g^{i,k})\right\}^{\hat{x}_g^{i,k}} \times \left\{1-L_g^{i,k}(\epsilon^k|x_g^{i,k})\right\}^{4-\hat{x}_g^{i,k}}\right]^{z^{i,k}},
\end{aligned}$$

where $\epsilon=(\epsilon^k)_{k\in K}$, $z^i=(z^{i,k})_{k\in K}$, and $x^i=(x_g^{i,k})_{k\in K}^{g=1,\cdots,4}$.

As a result, we can estimate the distribution of $z^i=(z^{i,k})_{k\in K}$ which allows us to categorize the subject i's individual mixing propensity: We categorize subject i to one of four types which has the highest $z^{i,k}$. Details follow.

[8]For illustration, suppose that a subject is a RO. With $\epsilon^{RO}=0$ or she does not make any mistakes, she will choose the action that maximizes the expected payoff with probability one. If $\epsilon^{RO}=1$ or she makes a choice in a completely random manner, then the probability of the optimal choice is $1/c_g$.

A1.1.1. Type Categorization

Each subject with a different mixing-propensity type has a different pattern for each game. Define $x^{i,k}_{g,j}$ is the number of subject i's decision that equals to the type k subject's decision in game g of set j. Similarly, $\hat{x}^{i,\,k}_{g}$ is the number of the sets of subject i's decision that equals to the type k subject's decision in game g. Consider an example subject i with play (UUUD, UUUM, UUDD, UUUU) in the game 1. Each entry of the vector corresponds to four actions played at each set. For the RO type subject, $x^{i\,RO}_{1} = (x^{i,\,RO}_{1,1}, \cdots, x^{i,\,RO}_{1,4}) = (3,$ 3, 2, 4) and $\hat{x}^{i,\,RO}_{1} = 1$. Similarly, for the PM type, $x^{i\,PM}_{1} = (4, 3, 3, 3)$ and $\hat{x}^{i,\,RO}_{1} = 1$. In set 2, the subject played M instead of D (what was supposed to play as the PM type), we count three U actions as the matching actions for the PM type and D as a mismatching action. Similarly, in set 3 we count the excess number of D action as the mismatching action since the subject played D more than one time. The UM-type subject is supposed to play either U or D at each set, so $x^{i\,UM}_{1} = (3, 3, 2, 4)$ gives the highest matching entries. At the set 1, 2, and 4, the action U is interpreted as a dominant action and at the set 3, D is interpreted as the dominant action, so $\hat{x}^{i,\,UM}_{1} = 1$. For the HM type, $x^{i\,HM}_{1} = (0, 1, 0, 0)$ and $\hat{x}^{i,\,HM}_{1}$ $= 0$. Table 11 shows examples of choice patterns of the RO-type subject and the PM-type subject.

Rational Optimizer				
Set 1-4	Game1	Game2	Game3	Game4
R1	U	U	ALL	ALL
R2	U	U	ALL	ALL
R3	U	U	ALL	ALL
R4	U	U	ALL	ALL

Probability Matcher				
Set 1-4	Game1	Game2	Game3	Game4
R1	U	U	U	U
R2	U	D	U	MU
R3	U	U	M	MD
R4	D	D	D	D

Table 11: Examples of the RO-type (Left) and the PM-type (Right) subject's choice patterns

In Games 1 and 3, the RO-type subject is supposed to play U for every round of every set. In Games 2 and 4, any choice will be accepted as the optimal choice since all actions are expected to give the exactly the same payoff. We identify the RO-type from the other types by the pattern of plays exhibited in Game 1 and 3.

We can distinguish the PM-type subject from the other type by observing their mixing proportion of plays: The proportion must be kept across the sets and matched to the given distribution of the virtual players' type. In Game 1, the PM-type subject is supposed to play U three times and D once. The order of play is irrelevant as long as the frequency is kept to 3 Us and 1 D in every set. In Game 2, U and D is supposed to be played two times respectively in every set. This proportion of mixing play is matched to the given distribution of the virtual player's type (L: R = 1: 1). In Game 3, to match the given distribution of the virtual player's type, the action U is supposed to be chosen two times, M and D is supposed to be chosen one time each in every set. In Game 4, all actions (U, MU, MD, and D) are supposed to be chosen one time each in every set.

Uniform Matcher					Hedging Matcher				
	Game1	Game2	Game3	Game4		Game1	Game2	Game3	Game4
Set 1	U4	U4	U4	U4	Set 1	M4	M4	B4	B4
Set 2	U4	U4	U4	MU4	Set 2	M4	M4	B4	B4
Set 3	U4	D4	M4	MD4	Set 3	M4	M4	B4	B4
Set 4	D4	D4	D4	D4	Set 4	M4	M4	B4	B4

Table 12: Examples of the UM-type (Left) and the HM-type (Right) subjects' choice patterns

Table 12 shows examples of choice patterns of the UM-type subject and the HM-type subject. The UM-type subject is supposed to play the same action within each set. Different to the RO type, they may change their action in each set so that they match their entire choices' mixing proportion to the probability distribution of the virtual player's type. For example, in Game 1, the UM-type subject may choose four Us in three sets and four Ds in the other set. In Game 2, the UM-type subject may choose four Us in two sets and four Ds in the other two sets. In Game 3, U would be played in two sets, and M and D would be chosen in one set each. In Game 4, each one action would be played in each set. By observing such choice patterns, we can distinguish the UM-type subject from the other type.

The HM-type subject is supposed to play the hedging actions which always provide some positive payoff. In each game, such hedging action (M in Games 1 and 2, and B in Games 3 and 4) would provide a discounted payoff according to their own risk preferences, so it cannot be beneficial than playing the rational action. Since the HM-type subject act in a homogeneous manner, we can distinguish the HM-type easily from the other types.

A1.2. SDM Model Specification

A1.2.1 Information and Payoff

In each round, subjects face the set of information $([a^1, b^1], p^1; [a^2, b^2], p^2)$, where $[a^i, b^i]$ is the range within which player $i \in \{1, 2\}$ can choose a number, and p^i is the target parameter for player i. That is, each subject completely knows his own and the partner's strategic environments. In every game, subjects are notified that they play the role of player 1 and their partner plays the role of player 2. We denote $x^i \in [a^1, b^1]$ as a choice of player i. In each round, player 1 earns payoffs which depend on his choice x^1 and player 2's choice x^2. We define the payoff function $P(x^1|x^2; [a^1, b^1], p^1, [a^2, b^2], p^2)$ as follows:

$$P(x^1|x^2; [a^1, b^1], p^1, [a^2, b^2], p^2) = 100 \left[1 - \frac{\left| \ln\left(\frac{x^1 - p^1 \cdot x^2}{|a^1 - p^1 \cdot b^2|} + 1 \right) \right|}{\ln\left(b^1 - a^1 + p^1(b^2 - a^2)\right)} \right]$$
$$\equiv 100 \left[1 - \frac{\left| \ln\left(\frac{x^1 - p^1 \cdot x^2}{|\underline{e}^1|} + 1 \right) \right|}{\ln\left(\overline{e}^1 - \underline{e}^1\right)} \right],$$

where $\overline{e}^1$ and $\underline{e}^1$ denote the largest and smallest possible differences between x^1 and $p^1 \cdot x^2$ respectively. That is, $\overline{e}^1 \equiv b^1 - p^1 \cdot a^2$ and $\underline{e}_1 \equiv a^1 - p^1 \cdot b^2$. For notational simplicity, we denote $P(x^1|x^2; [a^1, b^1], p^1, [a^2, b^2], p^2) \equiv P^1(x^1|x^2; e^1)$, where $e^1 = (\overline{e}^1, \underline{e}^1)$.

Player 1 can maximize his own payoff by minimizing the difference between x^1 and $p^1 \cdot x^2$ with respect to his own prediction of x^2. Suppose that player 1 believes that his partner chooses a certain action x^2 in an interval $[\underline{x}^2, \overline{x}^2] \subseteq [a^2, b^2]$ and each x^2 is uniformly distributed within the interval. We denote such a belief as a probability distribution $f^1(x^2|\underline{x}^2, \overline{x}^2)$. Then, the expected payoff from player 1's choice x^{1*} is given by

$$E[P^1(x^{1*}|x^2; e^1)] = \int_{\underline{x}^2}^{\overline{x}^2} \left[100 \left(1 - \frac{\left| \ln\left(\frac{x^{1*} - p^1 \cdot x^2}{|\underline{e}^1|} + 1 \right) \right|}{\ln(\overline{e}^1 - \underline{e}^1)} \right) \right] f^1(x^2|\underline{x}^2, \overline{x}^2) dx^2.$$

Then, the optimal choice x^{1*} that maximizes $P^1(x^{1*}|x^2; e^1)$ satisfies the following equation:

$$(x^{1*} - p^1 \cdot \underline{x}^2 + |\underline{e}^1|)(x^{1*} - p^1 \cdot \overline{x}^2 + |\underline{e}^1|) = (|\underline{e}^1|)^2$$

There are two notable features about the payoff function worth mentioning. First, due to the asymmetric concavity of the payoff function, it distinguishes a point-based prediction with an interval-based prediction. That is, the payoff decreases in a concave manner when $|x^1 - p^1 \cdot x^2| > 0$ and the slope of function when $x^1 - p^1 \cdot x^2 > 0$ is different to that when $x^1 - p^1 \cdot x^2 < 0$. This modification is required to distinguish the subject who uses an exact number as a prediction from the one who uses the interval that may have the same mean under the uniform distribution. For example, suppose that player 1 has $p^1 = 1.5$ and

$[a^1, b^1] = [100, 900]$. Consider two cases in which (1) player 1 predicts player 2 plays exactly 300 and (2) player 1 has a belief that player 2's choice is uniformly distributed within an interval $[100, 500]$. In the former case, player 1 may choose 450 to capture $p^1 \cdot x^1 = 1.5 \times 300$. In the latter case, player 1 may choose $50(\sqrt{205} - 4) \simeq 515.89$ to best-respond to his own belief. Costa-Gomes and Crawford (2006) (henceforth CGC) adopted a kinked linear function that imposes different linear slopes at two different intervals. Even though they avoided a simple linear function, they couldn't distinguish the choice from the point-based prediction and that from the interval-based prediction. In CGC, two different predictions will lead to the same choice of 450. [9] Imposing an asymmetric concavity on the payoff function can be useful device to avoid those two belief structures. Second, we normalize the payoff function by using the difference of the largest possible prediction $\overline{e}_1$ and the smallest possible prediction $\underline{e}_1$. This normalization helps subjects attain similar amount of payoffs in every game. Since each game has a different choice interval, the extent of the prediction error also can vary by games. The normalization adjusts the unit of payoffs so that the extent of payoff loss from the prediction error will be measured in a less heterogeneous manner.

A1.2.2 Statistical Model Specification

In the SDM experiment, we focus on the identification of the individual strategy with respect to the mixing propensity, which requires the estimation of an individual strategy and the mixing propensity type respectively. For this reason, the estimation was conducted by a two-layer process. In the first layer, we fix the individual mixing propensity k among four types (RO, PM, UM, and HM). Then, given fixed individual type, we guess a type-specific strategy $s^i(k)$ of subject i with respect to such fixed type. We allows multiple type-specific strategies. In the next subsection we discuss how we can guess the type-specific strategies from the actual data. Having such a set of strategies $S^i(k)$, we use the maximum likelihood method to estimate the probability distribution of likelihood for each strategy. From the result of the estimation, we pick the most probable type-specific strategy $s^{i*}(k)$ for type k. In the second layer of the estimation, we collect the most probable type-specific strategies for each type k and define such set of four type-specified strategies

[9] Distinguishing the point-based from the interval-based prediction is a sensitive issue as long as we are based on the iterated dominance model. In either Lk or CH model, the L1 strategy will be based on the interval-based prediction rather than the point-based prediction. That is, both models assume that L1 strategy users believe that the L0 player plays randomly over the interval. On the other hand, the point-based prediction is different in that the arbitrary belief anchors on a certain point in the interval. For that reason, separating those two cases will provide a useful evidence to confirm that individuals develop their prediction based on the belief that L0 players exist.

$S^{i*} \equiv \{s^{i*}(RO), s^{i*}(PM), s^{i*}(UM), s^{i*}(HM)\}$ as a set of types for the second layer. Among those four types of S^i, we estimate the most probable type by using the maximum likelihood method.

$a^i_{j,l}$ and $b^i_{j,l}$ denote subject i's lower and upper bound in round l of set j respectively. $x^i_{j,l}$ is subject i's unadjusted guess in round l of set j. Considering that $x^i_{j,l}$ may be constrained by the bounds, we define an adjusted guess $R(x^i_{j,l}) \equiv \min\{b^i_{j,l}, \max\{a^i_{j,l}, x^i_{j,l}\}\}$ which restricts the actual choice into the interior of the bounds in each round. We define a target guess for an individual with type s in round l of set j as $t^s_{j,l}$. Since our experiments allow subjects to choose integer values only, $T^{i,s}_{j,l} \equiv [t^s_{j,l} - 0.5, t^s_{j,l} + 0.5] \cap [a^i_{j,l}, b^i_{j,l}]$ is a target bound for the adjusted guess of individual i with type s in round l of set j. That is, $T^{i,s}_{j,l}$ restricts choose-able bound around the exact target $t^s_{j,l}$ with respect to a small possibility of error (± 0.5). Since we assume that all subjects can find the correct guess by using the calculation panel, we only allow very narrow range around the exact target guess.

$\epsilon^s \in [0,1]$ is a type-specific error rate of the adjusted guess and $d^s(R(x^i_{j,l}), \lambda)$ is an error density of an individual with type s, with a precision level λ for the adjusted guess in round l of set j. For precision level λ, we assume λ to be the same across the sets and rounds. We assume that ϵ^s is identically and independently distributed over all the rounds.

$P^i_{j,l}(x|y)$ is subject i's payoff from his own guess x given his partner's guess y in round l of set j. From this payoff, we define the expected payoff of an individual with type s in round l of set j as $P^{i,s}_{j,l}(x)$:

$$P^{i,s}_{j,l}(x) \equiv \int_{a^i_{j,l}}^{b^i_{j,l}} P^i_{j,l}(x|y) f^s_{j,l}(y) dy$$

where $f^s_{j,l}(y)$ is a density of y that is distributed according to the belief of type s.

We assume that a "spike-logit" shape of error.[10] With this assumption, $d^s(R(x^i_{j,l}), \lambda)$ is defined as :

$$d^s(R(x^i_{j,l}), \lambda) = \begin{cases} \frac{\exp[\lambda P^{i,s}_{j,l}(R(x^i_{j,l}))]}{\int_{[a^i_{j,l}, b_{j,l}] \setminus T^{i,s}_{j,l}} \exp[\lambda P^{i,s}_{j,l}(z)] dz} & \text{for } R(x^i_{j,l}) \in [a^i_{j,l}, b_{j,l}] \setminus T^{i,s}_{j,l} \\ 0 & \text{for } R(x^i_{j,l}) \in T^{i,s}_{j,l}. \end{cases}$$

We define $n^{i,s}_j$ the number of rounds that subject i with type s plays the exact guess

[10] We assumed the distribution of error with the consideration that the error rate to be decreased with convex rate. [THIS IS GOT TO BE ASKED!! HEECHUN WHAT DID YOU TRY TO SAY HERE?]The use of calculation panel reduces the possibility of error that purely comes from miscalculation. Moreover, with the consideration of rounding, we allowed the range of exact choice to include the closest integer.

of type s in set j and $N_j^{i,s}$ a collection of such rounds in set j. We define vectors $x_j^i \equiv (x_{j,1}^i, x_{j,2}^i, \cdots, x_{j,5}^i)$ as the guesses of subject i, and $R(x_j^i) \equiv (R(x_{j,1}^i), R(x_{j,2}^i), \cdots, R(x_{j,5}^i))$ as the adjusted guesses of subject i in set j.

Considering that individual i with type s chooses the (adjusted) guess $R(x_{j,l}^i)$ with probability $1-\epsilon^s$, we have a sample density for $R(x_{j,l}^i)$ in set j, denoted by $d^s(R(x_j^i), \epsilon^s, \lambda)$ as follows:

$$d^s(R(x_j^i), \epsilon^s, \lambda) \equiv (1-\epsilon^s)^{n_j^{i,s}} (\epsilon^k)^{5-n_j^{i,s}} \prod_{l \notin N_j^{i,s}} d^s(R(x_{j,l}^i), \lambda).$$

Similarly, we define $R(x^i) \equiv (R(x_1^i), R(x_2^i), \cdots, R(x_8^i))$ as subject i's adjusted guesses for the entire experiment and $d^s(R(x^i), \epsilon^s, \lambda)$ as a sample density function for the entire experiment :

$$d^s(R(x^i), \epsilon^s, \lambda) \equiv \prod_{j=1}^{8} d^s(R(x_j^i), \epsilon^s, \lambda).$$

Now we define $z_j^{i,s}$ as an indicator of type s for subject i, where $z_j^{i,s} = 1$ if the subject is of type s and $\sum_{s \in S} z_j^{i,s} = 1$. $\epsilon \equiv (\epsilon^s)_{s \in S}$ is a vector of error rates for all types and $z_j^i \equiv (z_j^{i,s})_{s \in S}$ is a vector of type indicators in set j. From these definitions, we have subject i's log-likelihood function $L(z_j^i, \epsilon, \lambda | R(x_{j,l}^i))$ as follows:

$$L(z_j^i, \epsilon, \lambda | R(x_{j,l}^i)) \equiv \sum_{s \in S} z_j^{i,s} \ln \left[d^s(R(x^i), \epsilon^s, \lambda) \right].$$

A1.2.3. Type Categorization

We use actual choice data of the SDM experiment to guess the possible set of specific strategies per type from the ODM experiment. Unlike the ODM experiment, the SDM experiment does not provide explicit ways of belief formation that individuals are expected to follow. The RO-type subjects, who might use a single action for the whole experiment, are relatively easy to identify. On the other hand, identifying the other types, PM, UM, and HM types, requires us to find not only strategies that subjects use, but also the mixing proportion among those strategies. This consideration enforces us to (theoretically) try the infinite number of different mixing proportions with different strategies. For example, a PM-type subject who adopts L1 strategy and L2 strategy with mixing proportion 0.75 and 0.25 and another PM-type subject who adopts L1 and L2 strategies with mixing proportion

0.50 and 0.50 should be classified as different types. To avoid such difficulties, we restrict our attention to a limited set of strategies. To this end, we exploit the actual choice observations to guess the probable strategies and the mixing proportion among them by individuals.

- Rational Optimizer (RO) type

RO-type subjects always have a fixed set of type-specific strategies. That is, RO-L1-type subject is expected to always choose L1 strategy, RO-L2-type subject is expected to always choose L2 strategy, and so on. Since we restrict our attention to only four strategies (L1, L2, L3 and NE), this assumption restricts the set of the RO-type subjects' strategies. That is, we construct the strategy set for RO-type subject i as $S^i(RO) = \{RO-L1, RO-L2, RO-L3, RO-NE\}$. From four strategies, we find the specific strategy $s^{i*}(RO) \in S^i(RO)$ that maximizes the likelihood among them.

Game	RO–L1	RO–L2	RO–L3	RO–NE
$\alpha n_2 \beta n_4$	419.4	361.1	440.3	500
$\beta n_4 \alpha n_2$	515.9	629	541.8	750
$\delta n_3 \beta n_1$	678.3	363.9	373.1	300
$\beta n_1 \delta n_3$	330.8	339.2	181.9	150
$\beta n_1 \beta n_2$	350	173.9	122.5	100
$\beta n_2 \beta n_1$	347.8	245	121.75	100
$\delta n_3 \gamma n_3$	300	550	363	550
$\gamma n_3 \delta n_3$	500	330	500	500

Table 13: Predicted patterns of play in the SDM experiment for RO-type subjects

In each round of the sets, the RO-type subject is supposed to play a certain action which corresponds to the optimal strategy. For example, the RO–NE-type subject would play the NE strategy in each round. While the most of the rounds allow the distinction of different strategies, Games $\delta n_3 \gamma n_3$ and $\gamma n_3 \delta n_3$ share the same numbers for different strategies. In Game $\delta n_3 \gamma n_3$, L2 player and NE player can play the same choice 550, and in Game $\gamma n_3 \delta n_3$, the strategy type for a subject who chose 500 would not be distinguished. To distinguish them, we need to rely on the records from the calculation panel. In Game $\delta n_3 \gamma n_3$, L2 player would use the calculation panel to calculate L1 partner's choice, 500, and then by putting 500 into his own calculation panel he would learn that 550 is a best-responding choice. NE player, to get 550, may start with calculating his own choice for L1 strategy and have 300 from the initial calculation. Also, NE player may repeatedly use (more than 3 times) the calculation panel to arrive the NE strategy.

(2) Probability Matcher (PM) type

For identifying PM-type subjects' strategies, we need to specify not only strategies but also the mixing proportion among the strategies. In case of the RO type, picking a certain strategy from the fixed set is enough for identification of the type-specific strategy. However, the PM-type subjects may use more than one strategy with respect to their own belief structure. For this reason, we need to specify the multiple strategies they may adopt and the frequency how often the strategies are used together. The way how we can guess them simultaneously is the main concern for identification of PM-type subjects' strategies. What we did is as follows: For identifying (pure) strategies, we restrict our focus to 11 different combinations of strategies. Since we assume that subjects use only four pure strategies (L1, L2, L3, and NE), PM-type subjects could have (i) 6 different combinations of 2-strategy cases: L1 + L2, L1+ L3, L1 + NE, L2 + L3, L2 + NE, and L3 + NE, (ii) 4 different combinations of 3-strategy cases: L1 + L2 + L2, L1 + L2 + NE, L2 + L3 + NE, L1 + L3 + NE, and (iii) one combination of 4-strategy case. For illustration, we consider a L1 + L2 case described in Table 14. That is, in n this example we describe hypothetical choices of a PM-type subject who uses L1 and L2 strategies and mixes them with proportion L1 : L2 = 3 : 2. Each two columns describe the actions of each round and corresponding strategies. The first column shows choices of the PM-type subject of L1 + L2 strategy. The subject may use either pure L1 strategy or L2 strategy. The next column shows the corresponding strategies for each choice. For example, subject's action 419.4 (of the first column) in Game $\alpha n_2 \beta n_4$ of round 1 corresponds to L1 strategy (of the second column).

We first identify which strategies are used in each set and then find an average proportion among them. In this process, we exclude choices that does not have corresponding strategies from L1, L2, L3, and NE. Once we have an average proportion among those strategies, we round up/down it to be fitted with 5-round setting. Similarly, a PM-type subject who uses different collection of strategies with different mixing proportion can be classified. This process is based on the assumption that PM-type subjects would keep the same mixing proportion across the sets for the same collection of pure strategies. This assumption allows us to guess the mixing proportion from an observed average proportion of choices. From a vector $x^i \equiv (x^i_1, x^i_2, \cdots, x^i_5)$, we can find the frequency of each choice that corresponds to each

Probability Matcher with L1(60%)+L2(40%)										
Game	Round1	Stg.	Round2	Stg.	Round3	Stg.	Round4	Stg.	Round5	Stg.
$\alpha n_2\beta n_4$	419.4	L1	361.1	L2	419.4	L1	361.1	L2	419.4	L1
$\beta n_4\alpha n_2$	515.9	L1	515.9	L1	515.9	L1	521.7	L2	521.7	L2
$\delta n_3\beta n_1$	678.3	L1	363.9	L2	363.9	L2	678.3	L1	678.3	L1
$\beta n_1\delta n_3$	339.2	L2	330.8	L1	330.8	L1	330.8	L1	339.2	L2
$\beta n_1\beta n_2$	350	L1	173.9	L2	173.9	L2	350	L1	350	L1
$\beta n_2\beta n_1$	245	L2	347.8	L1	245	L2	347.8	L1	347.8	L1
$\delta n_3\gamma n_3$	550	L2	300	L1	300	L1	300	L1	550	L2
$\gamma n_3\delta n_3$	330	L2	500	L1	500	L1	330	L2	500	L1

Table 14: An example of a PM-type subject's pattern of play in the SDM experiment

strategy. Then, we use the the observed frequency as the guess for the mixing proportion.

From this process, we find at most five candidates of strategies specified for PM-type subjects from each set-wise observations. For example, consider subject i's choices in set 1 that shows the proportion among each strategy as L1 : L2 = 2 : 1 : 1 : 1. We name a type-specified strategy which shows the mixing proportion L1 : L2 : L3 : NE = 2 : 1 : 1 : 1 as PM-1. Then, we compare the actual choices in other sets (set 2–set 8) with this PM-1 strategy. Similarly the actual choices in set 2 shows L1 : L2 : L3 : NE = 1 : 2 : 1 : 1. Then, we have another type-specified strategy PM-2 with the mixing proportion L1 : L2 : L3 : NE = 1 : 2 : 1 : 1, so on. As a result, we have at most eight different guesses of mixing strategies specified for PM-type subjects.

(3) Uniform Matcher (UM) type

For the UM type, we consider the observations from the entire experiment to guess the type-specific strategy. Different to what we did for the PM type, we assume that UM-type subjects play the same strategy within each set but the strategy will be changed across the sets. For this reason, we could easily guess the strategies used in each set. To guess the mixing proportion, we consider the frequency among those strategies from the entire experiment. That is, we only count the all the choices played in the experiment (from set 1 to set 5, exclude all non-RO-type actions). Naturally, the actual mixing proportion is the guess for the mixing proportion. As an example, we consider a UM-type subject with strategies L1+L2+NE with mixing proportion L1 : L2 : NE = 4: 2: 2. In Table 15, the UM-type subject uniformly played L1 strategy in set 1, L2 strategy in set 2 and 4, and NE strategy in set 3 and 5. From this observation, we can guess the collection of strategy adopted by the subject as L1, L2 and NE. The mixing proportion among the adopted strategies can be guessed by counting the mixing proportion in the entire experiment.

Uniform Matcher with L1(50%)+L2(25%)+NE(25%)										
Game	Round1	Stg.	Round2	Stg.	Round3	Stg.	Round4	Stg.	Round5	Stg.
$\alpha n_2 \beta n_4$	419.4	L1	419.4	L1	419.4	L1	419.4	L1	419.4	L1
$\beta n_4 \alpha n_2$	515.9	L1	515.9	L1	515.9	L1	515.9	L1	515.9	L1
$\delta n_3 \beta n_1$	300	NE	300	NE	300	NE	300	NE	300	NE
$\beta n_1 \delta n_3$	339.2	L2	339.2	L2	339.2	L2	339.2	L2	339.2	L2
$\beta n_1 \beta n_2$	173.9	L2	173.9	L2	173.9	L2	173.9	L2	173.9	L2
$\beta n_2 \beta n_1$	347.8	L1	347.8	L1	347.8	L1	347.8	L1	347.8	L1
$\delta n_3 \gamma n_3$	550	NE	550	NE	550	NE	550	NE	550	NE
$\gamma n_3 \delta n_3$	500	L1	500	L1	500	L1	500	L1	500	L1

Table 15: An example of a UM-type subject's pattern of play in the SDM experiment

For the actual guess of the possible set of specific strategies for the UM type, consider a UM-type subject with observed strategies L1 : L2 : L3 : NE = 15 : 10 : 13 : 2 as another example. From this observation, we can find the relative proportion among the strategies. However, we can imagine the situation in which the observed frequency among the strategies may not clearly fit to 8-set setting, or multiple strategies appear within the same set. For the former case, we rounded up or down for the proportions to be the multiples of 1/8. For example, when we round up or down the relative proportion L1 : L2 : L3 : NE = 15 : 10 : 13 : 2 = 0.375 : 0.25 : 0.325 : 0.05, we could have several different approximations. That is, 0.375 : 0.25 : 0.325 : 0.05 can be rounded to 0.375 : 0.25 : 0.375 : 0.125, or to 0.375 : 0.25 : 0.375 : 0. When we have such several different approximations, we consider the most frequently used strategy in each set as a major strategy for that set. For the case that rounds up 0.325 to 0.375 = 3/8 for L1, three sets are required to uniformly play L1. Then we find three sets that L1 appears mostly. For such sets, we count non-L1 choices as deviations from L1 strategy. We can apply this process to all the other strategies and derive the sample density $d^{UM}(R(x^i), \epsilon^{UM}, \lambda)$.

(4) Hedging Matcher (HM) type

For HM type, we allow HM-type subjects to choose non-Lk choices and this relaxation leads us to another difficulty; whether to consider such choices as a strategic hedging behavior or not. For this concern, we exploit two assumptions: (1) Any hedging behavior would be based on the belief that consists of multiple Lk strategies, and (2) any hedging behavior based on a certain belief would be bounded by the interval formed from the belief. For example, consider a HM-type subject in game $\alpha n_2 \beta n_4$, who has a belief that his/her partner may play either L1 or L2 strategy with some mixing proportion. Then, his/her belief may form a bound for his/her hedging choice and that bound may depend on the L1 (419.4) and L2 (361.1) strategies. Given his/her belief of L1 and L2, choosing any numbers outside of the interval formed by 419.4 and 361.1 (in this case, [361.1, 419.4]) is always weakly dominated strategy by any number in the interior of the interval. From these assumptions, we can infer that any HM-type subjects may choose the number within the interval that is bounded by Lk strategies he/she based on. This inference, even though it allows broader range than that for UM- or PM-type subjects, provides ground to identify whether the subject shows consistent HM-type behavior. As an example, consider an HM-type subject with strategies L1+L2. Table 16 shows that all choices taken by the HM-type subject are consistently located in the interval of L1 and L2 strategies. This pattern of play can distinguish the HM-type subject from random-playing subjects.

Hedging Matcher with L1+L2							
Game	Round1	Round2	Round3	Round4	Round5	L1	L2
$\alpha n_2 \alpha n_4$	380	400	415	375	365	419.4	361.1
$\alpha n_4 \alpha n_2$	516	580	600	550	629	515.9	629
$\alpha n_4 \beta n_1$	400	450	650	550	600	678.3	363.9
$\beta n_1 \alpha n_4$	332	333	333	338	335	330.8	339.2
$\beta n_1 \beta n_2$	350	180	250	200	300	350	173.9
$\beta n_2 \beta n_1$	250	300	333	325	280	347.8	245
$\delta n_3 \gamma n_3$	300	550	350	500	400	300	550
$\gamma n_3 \delta n_3$	350	400	500	450	400	500	330

Table 16: An example of a HM-type subject's pattern of play in the SDM experiment

For HM type, specifying the adopted strategies would be sufficient. Different from what we did for UM or PM type, we focus on identifying HM-type subjects itself. We consider 6 different combinations of two strategies that result in different intervals: L1 + L2, L1 + L3, L1 + NE, L2 + L3, L2 + NE, and L3 + NE. Even though the range of target guess would be wide, we can effectively identify the consistent behavior from the HM type. First, every combination of strategies has the different formation which depends on the game structure. For example, L1 + NE may have the broadest range in most of games but not for game 1 and 8. Therefore, even if a subject who behaves randomly may seem to consistently choose the number within the interval with L1 and NE strategies, such a subject may show consistent deviation in every round of game 1 and 8. Moreover, the shape of the interval in the game also changes. In games 1, 2 and 7, the choices from L1 strategy are lower than that from NE strategy, while the opposite is true in games 3, 4, 5, and 6. From this structure, HM-type subjects with a certain belief need to know the exact range that will bound his/her optimal hedging choice. Moreover, the ranges that are covered by such combinations are at most equal to or smaller than a half of the entire choice interval. For this reason, the probability that the subject who consistently chooses numbers bounded by random behaviors is practically negligible.

Appendix B. Experiment Instructions

B1. The ODM Experiment

Preliminary Survey:

Before you start the actual task, we'd like to ask you to answer three survey questions. Consider a hypothetical situation where I give you either a fixed amount of money, X, or a simple lottery ticket. You can choose only one. The lottery ticket, denoted by (W, p), gives you money W with probability p, or nothing with probability $1-p$. For example, $(\$100, 0.4)$ is a lottery ticket that gives you \$100 with probability 0.4, or nothing with probability 0.6. We want you to compare such lottery tickets with fixed amounts of money. [Three questions asking certainty equivalences of (\$10,0.5), (\$1,0.3), and (\$1,0.7), respectively.]

Important Preliminary: "... randomly drawn from a probability distribution??"

We want you to understand what we mean by "an event is randomly drawn from a probability distribution." A probability distribution is a description of possible events and their chances. For example, if you toss a fair coin, with 50% of chance it will land heads (H) and with another 50% of chance it will land tails (T). Here the possible events are the faces of a coin, H and T, and the corresponding chances are 0.5 each. Then the probability distribution of coin tossing can be described as (H,0.5; T,0.5). When we say "an event (here, the face of a coin) is randomly drawn from (H,0.5; T,0.5)," it means that we toss a coin and either H or T is realized, but we will not tell you what the actual realization is.

Here is another example. If we say "an event is randomly drawn from (L,0.2; C,0.5; R,0.3)," then it means three things: (1) An event L (,C or R) will be drawn with 20% (, 50% or 30%) of chance. (2) One among L, C and R is realized according to their chances. (3) We will not tell you what the realization is.

During this experiment, you will frequently read "an event is randomly drawn from a probability distribution" in various contexts. We will assume now that you completely understand the meaning of the sentence. Please raise your hand if you need further explanation.

Description of Games:

Your task is to make choices to earn points from several games as much as you can. One point is equivalent to 25 cents. The specific form of each game will vary, but it will be similarly described as the following payoff matrix.

	H, 0.3	T, 0.7
U	1	0
M	0.15	0.35
D	0	1

Your options will be shown on the left; U, M and D in this example. A probability distribution will be on the top; (H,0.3; T,0.7). The matrix shows your payoff. For example, if you choose M and event H is randomly drawn from the probability distribution, you will earn 0.15 points. If you choose D when event T is drawn, you will earn 1 point.

Here is another example. Suppose another game is described as the following matrix.

	L, 0.5	C, 0.3	R, 0.2
U	1	0	0
M	0	1	0
D	0	0	1
B	0.25	0.15	0.1

In this example, your options are U, M, D and B, and the event will be L with probability 0.5, C with probability 0.3, or R with probability 0.2. For example, if you choose B when event L is drawn, you will earn 0.25 points. If you choose U when event C is drawn, you will earn 0 points.

Structure of Experiment:

The experiment consists of 4 sessions. In each session you will play a game several times. Your available actions, a probability distribution over events, and payoffs will be informed.

Each session consists of 4 sets. In the beginning of each set, an event is randomly drawn from the probability distribution. Given this event, you will make choices for four rounds of the set. Note that the realized event will be unknown during those rounds. In the beginning of next set, another event is newly drawn from the same probability distribution. You will make another four choices, and so on.

Since a session consists of four sets and you have four rounds per set, one session runs 16 rounds. At a new session you will repeat the process with a new game which consists of new available actions, new probability distribution, and new payoffs. It is important for you to understand how a session and a set are defined, because in each session all 16 rounds look the same. For the first 4 rounds, i.e., the first set, the realized event is the same. For the next 4 rounds, i.e., the second set, an event is newly drawn and the realized event is the same.

Screening Test:

After you read the instructions, you will answer three multiple choice questions to check your understanding of the instructions. You may want to go back to review, or ask an experimenter for help. You can participate in the experiment only if ALL of your answers are correct. The main purpose of this screening test is to help you understand the instructions, not to cause you any stress. It is okay for you to ask an experimenter to help you if you are in doubt.

Q1. The experiment will consist of (A) sessions. In each session, you will have (B) sets. In each set, you will play (C) rounds of game. What are appropriate numbers in (A), (B) and (C)?

Q2. A game is described as the following matrix. [A matrix is displayed] Which of the followings is true?

Q3. Suppose a probability distribution over events is (L,0.3; C,0.5; R,0.2). Which of the followings is NOT true?

B2. The SDM Experiment

General description:

This experiment consists of 8 sets, and each set consists of 5 rounds of decision-making.

In each set, you are randomly matched to one experiment participant in this lab. You will play with your match for the set (5 rounds). In the next set, you will be matched to another participant. You will not know who your matches are, and your matches will not know you either.

In each round, you choose an integer in a certain range (e.g., between 100 and 300). You earn more when your number is closer to a certain target number (e.g., 0.7) times your match's actual choice. Your match will do the same task, but his/her range and target may be different.

Example:

Suppose that in the first set you choose a number between 0 and 300, your match chooses a number between 100 and 500, your target is 0.5, and your match's target is 1.1. If your match chooses, say, 200, then your payoff is maximized when you choose 100, because your target (0.5) times the match's choice (200) is 0.5*200 = 100. On the other hand, if your match believes that you will choose 100, then she would choose 110, because her target (1.1) times your choice (100) is 1.1*100 = 110.

Payoff:

Showing you the formula calculating the payoff is overkill: Because the components of the game (i.e., your range and target, and your match's range and target) change in every set, the formula contains some messy mathematical normalizations and adjustments. We instead, provide a calculator for you. Details will follow.

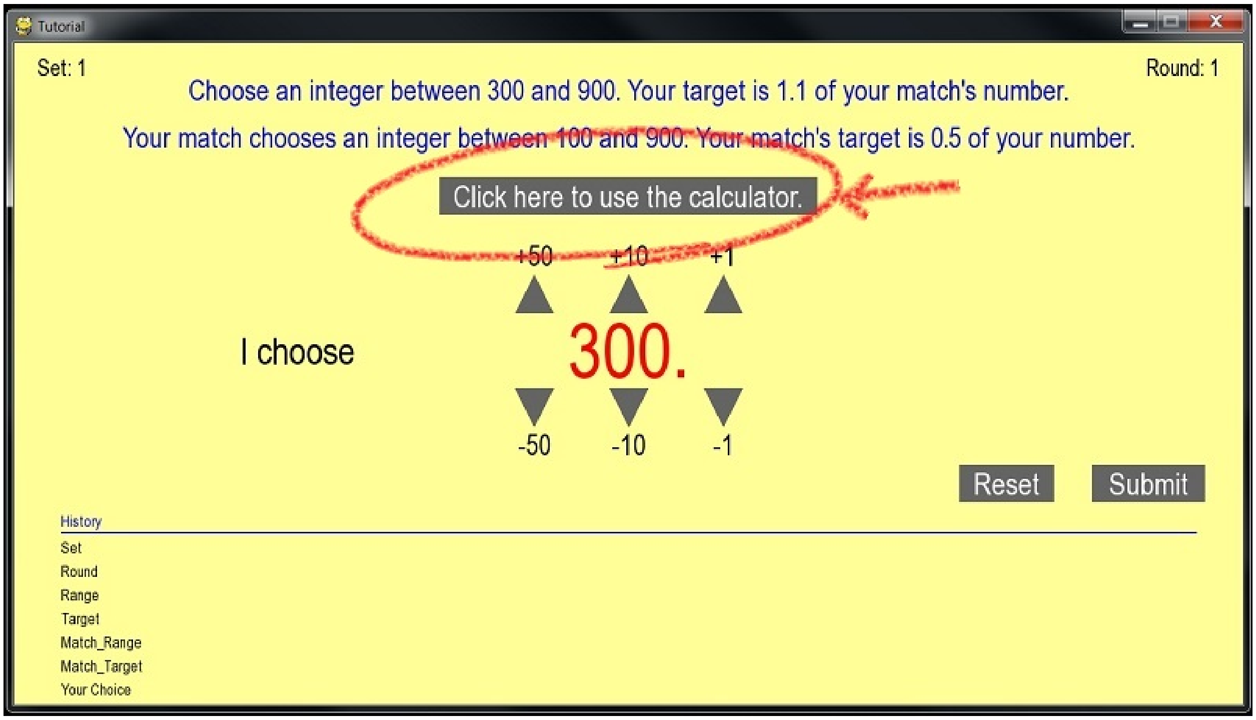
For now, please remember that your payoff is larger if you guess your match's choice more accurately. The calculation panel will do the remainder. If your choice is exactly the same as the 'ideal choice' (= your target * the match's choice), you will earn 25 cents for the round. The farther the difference between your choice and the ideal choice, the smaller you earn. The worst choice may give you 0 cents for the round.

Calculation Panel:

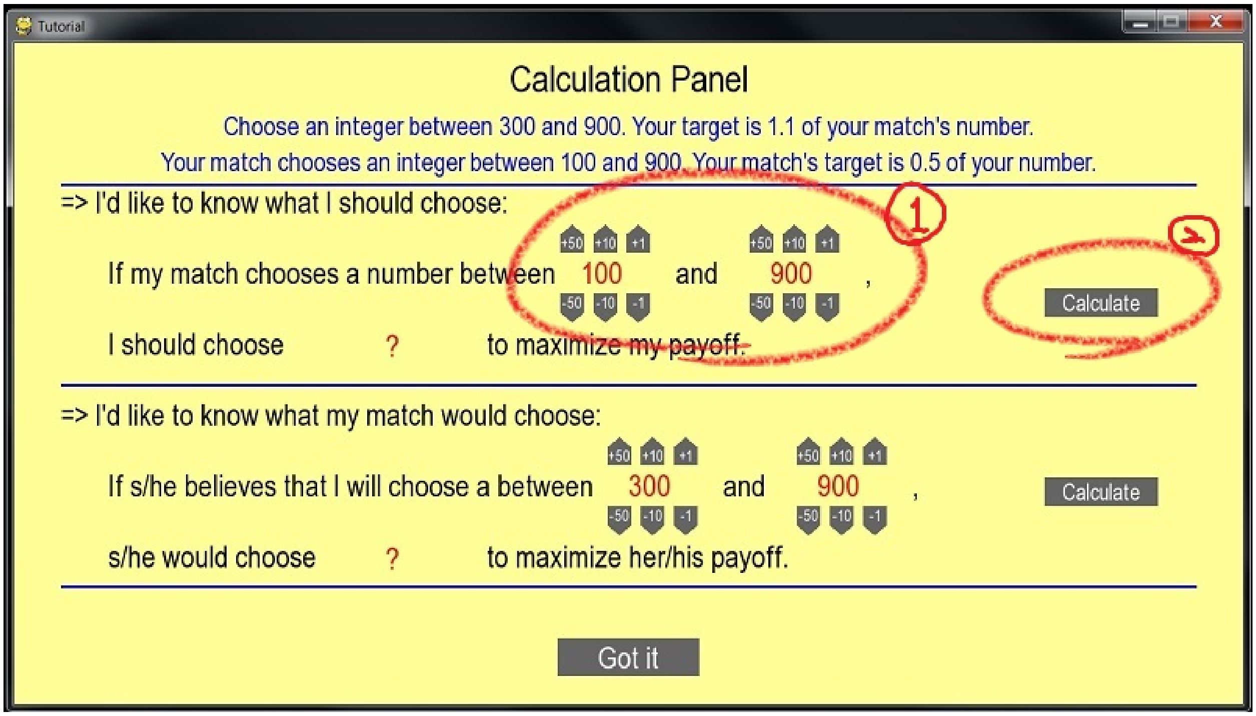
We want you to guess your match's choice, but do not want you to feel mathematically burdened. We provide a calculation panel. We highly recommend you to rely on the panel.

If you input your guess about your match's choice, then the calculation panel outputs your ideal choice. Also, if you input your choice, then the calculation panel outputs your match's ideal choice.

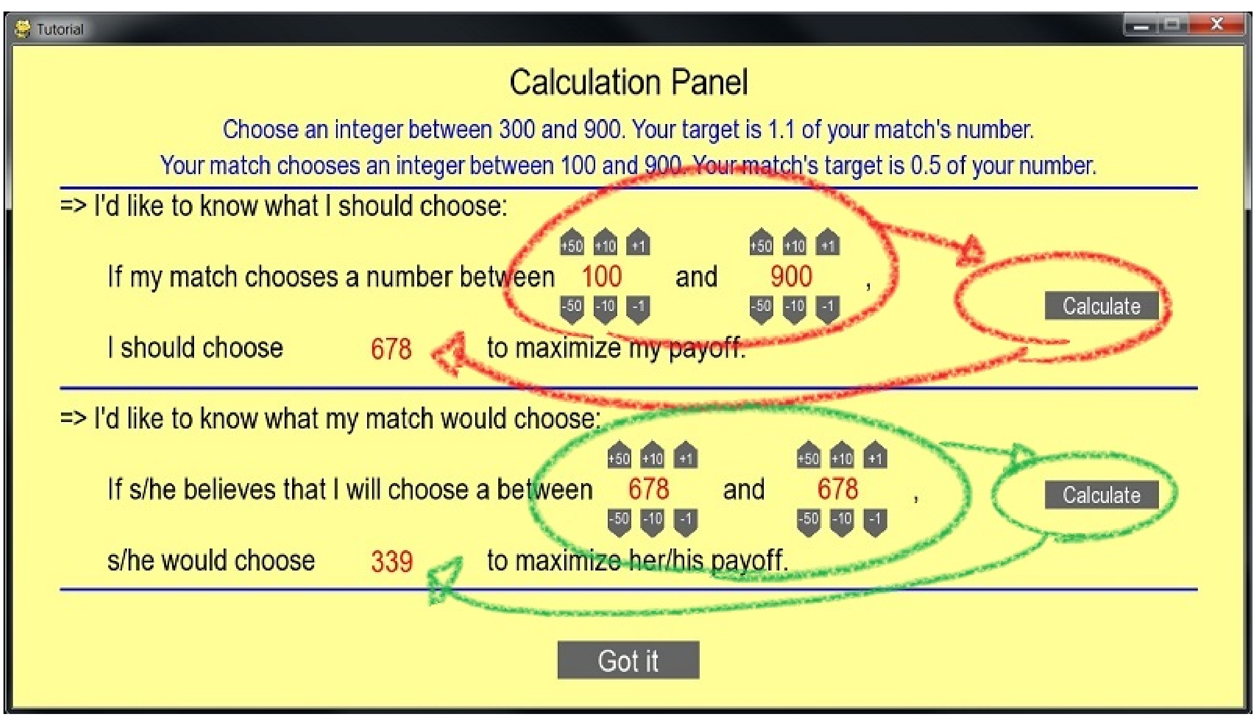
You can use the calculation panel as many times as you want. Check the following screenshots. You may also use paper and pencil to write down the past results of your calculation. The paper will be wasted anonymously. Your usage of calculation panel and paper will not affect your payoff.



Screenshot 1 [Set 1, Round 1]: To use the calculation panel, click the button indicated.



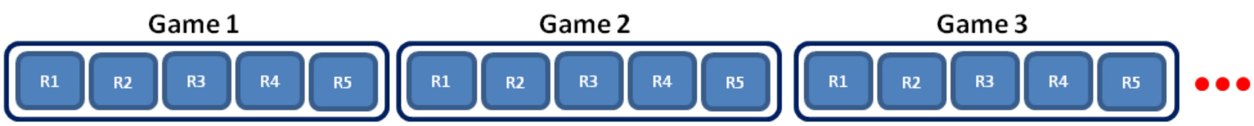
Screenshot 2 [Calculation Panel (1)]: If you want to know what you should choose, input your guess, and click 'Calculate' button. Similarly you can calculate what your match would choose.



Screenshot 3 [Calculation Panel (2)]: You can use this calculation panel as many times as you want. Note that it is okay to input the same number in the form of the range. (See a green circle above.)

Structure of the Experiment:

Again, this experiment consists of 8 sets. In the first set, you will play a game with a match for 5 rounds. In the next set, you will play another game with another match for five rounds, and so on. Even though the game will be the same for the five rounds within a set, we do not know whether your match will make the same choice for all five rounds.



* You play a game with a match for five rounds. After that, you will be matched with another participant.

Screening Test:

You will answer four multiple choice questions to check your understanding of this instruction. You may want to go back to review, or ask an experimenter for help. You can participate in the experiment only if ALL of your answers are correct. The main purpose of this screening test is to help you understand the instructions, not to cause you any stress. It is okay for you to ask an experimenter to help you if you are in doubt.

Q1. The experiment will consist of (A) sets. In each set, you will make (B) rounds of decision-making. What are appropriate numbers in (A) and (B)?

Q2. Suppose that your range is [100,500], your target is 0.5, your match's range is [100,900], and the match's target is 1.1. Which of the followings is NOT true?

Q3. You play a game with a match for five rounds. Which of the followings is true?

Q4. Suppose that you somehow guess that your match will pick any number between 200 and 500. Use the part of the calculation panel below and choose what you should choose. [The calculation panel is provided.]