CARF Working Paper

CARF-F-414

Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments and Theory

Yutaka Kayaba
University of Tokyo

Hitoshi Matsushima
University of Tokyo

Tomohisa Toyama
Kogakuin University

March 18, 2016
Current Version: July 14, 2017

CARF is presently supported by The Dai-ichi Life Insurance Company, Limited, Nomura Holdings, Inc., Sumitomo Mitsui Banking Corporation, Bank of Tokyo-Mitsubishi UFJ, Finatext Ltd., GCI Asset Management, Inc and The University of Tokyo Edge Capital Co.. This financial support enables us to issue CARF Working Papers.

Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments and Theory[1]

Yutaka Kayaba[2]
University of Tokyo

Hitoshi Matsushima[3]
University of Tokyo

Tomohisa Toyama[4]
Kogakuin University

March 18, 2016
Current Version: July 14, 2017

Abstract

We experimentally examine repeated prisoner's dilemma with random termination, in which monitoring is imperfect and private. Our estimation indicates that a significant proportion of subjects follow generous tit-for-tat strategies, straightforward extensions of tit-for-tat. However, the observed retaliating policies are inconsistent with the generous tit-for-tat equilibria. Contrary to the theory, subjects tend to retaliate more with high accuracy than with low accuracy. Specifically, they tend to retaliate more than the theory predicts with high accuracy, while they tend to retaliate lesser with low accuracy. In order to describe these results as unique equilibrium, we demonstrate an alternative theory that incorporates naïveté and reciprocity.

JEL Classification Numbers: C70, C71, C72, C73, D03.
Keywords: Repeated Prisoner's Dilemma, Imperfect Private Monitoring, Experiments, Generous Tit-for-Tat, Behavioral Model.

---

[1] Matsushima, Toyama, and Yagi (2007), Matsushima and Toyama (2011), and Matsushima, Tanaka, and Toyama (2013) are earlier versions of this paper. Research for this study was financially supported by a grant-in-aid for scientific research (KAKENHI 25285059) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese government, as well as by the Center of Advanced Research in Finance at the University of Tokyo. We are grateful to Hide Ichimura and the participants of seminars at SAET conference, Hitotsubashi University, and the University of Tokyo for their useful comments. We also appreciate the helpful support of Tomomi Tanaka and Nobuyuki Yagi. All errors are our own.
[2] Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email: ykayaba@e.u-tokyo.ac.jp
[3] Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hitoshi@e.u-tokyo.ac.jp
[4] Division of Liberal Arts, Kogakuin University, 1-24-2, Nishishinjuku, Shinjuku-ku, Tokyo 163-8677, Japan. E-mail: tomohisa.toyama@gmail.com

1. Introduction

It is a widely accepted that long-run strategic interaction facilitates collusion among players whose interests conflict with each other. The premise is that each player observes information about which actions opponents have selected before. However, even if the monitoring of the opponents' actions is *imperfect* (i.e., each player cannot directly observe the opponents' action choices but can observe informative signals), still the theoretical studies have shown that sufficiently patient players can employ cooperative strategies as equilibrium to a greater or lesser degree.

To be more precise, the folk theorem generally indicates that if the discount factor is sufficient, that is, close to unity, and each player can indirectly but not directly observe the opponents' action choices through noisy signals, a wide variety of allocations can be attained by subgame perfect equilibria in the infinitely repeated game (e.g., Fudenberg, Levine, and Maskin, 1994; Sugaya, 2012). Indeed, the folk theorem is applicable to a very wide range of strategic conflicts. However, the theorem does not specify what kind of equilibria emerge empirically, or the strategies people follow that are actually associated with the equilibria.

Given the lack of consensus on the strategies people empirically follow, this study experimentally analyses our subjects' behavior in the repeated prisoner's dilemma. Our experimental setup is imperfect monitoring. Each player cannot directly observe her opponent's action choice, but observes a signal instead, which is either *good* or *bad*. The good (bad) signal is more likely to occur when the opponent selects the cooperative (defective) action rather than the defective (cooperative) action. The probability that a player observes the good (bad) signal when the opponent selects the cooperative (defective) action is referred as monitoring accuracy, which is denoted by $p \in (1/2, 1)$. The study experimentally controls the levels of monitoring accuracy as treatments (high accuracy $p = 0.9$ and low accuracy $p = 0.6$) to examine the strategies our subjects follow. Specifically, the monitoring technology is *private* in that each player cannot receive any information about what the opponent observes about the player's choices (i.e., the signals are observable only to the receivers).

To examine the prevalence of the strategies our subjects employ, following the recent strand in the experimental repeated game literature, in which the heterogeneity of the strategies people follow is treated explicitly, we employ the strategy frequency estimation method (SFEM) developed by Dal Bó and Fréchette (2011). In the SFEM framework, we list various potential strategies of our subjects, and then estimate the frequencies of each strategy. This list includes strategies that share significant proportions in existing studies of experimental repeated games, such as tit-for-tat (TFT), grim-trigger, lenience, and long-term punishment. We include various non-Markovian strategies in the SFEM list.

Different to existing experimental studies (Fudenberg, Rand, and Dreber, 2012; Aoyagi, Bhaskar, and Fréchette, 2015), we rigorously include stochastic strategies in our SFEM list. Importantly, we include straightforward extensions of TFT, namely *generous tit-for-tat* (g-TFT).

Our SFEM estimates indicate that a significant proportion (about 70%) of our subjects follow g-TFT strategies, although they follow heterogeneous g-TFT strategies. G-TFT is a simple stochastic Markovian strategy that makes a player's action choice contingent only on the signal observed in the previous round and permits the stochastic choice between the cooperative action and the defective action in each round.[5] Because of the permission of stochastic action choices, g-TFT has a great advantage over the TFT strategy in equilibrium analysis: provided the discount factor is sufficient, g-TFT equilibria always exist, irrespective of the level of monitoring accuracy. Motivated by such theoretical importance, we regard g-TFT strategies and equilibria as the reasonable benchmarks of the standard repeated game theory with imperfect private monitoring. Our estimates imply that g-TFT is well supported empirically.

Observing that many of our subjects follow g-TFT strategies, we empirically examine their retaliation policies. We focus on the contrast of the probabilities of cooperative action choices contingent on a good signal and a bad signal, which is referred as the *retaliation intensity*.[6] Fixing a sufficient discount factor, the retaliation intensities are common across all g-TFT equilibria, depending on the level of monitoring accuracy. The retaliation intensity implied by the g-TFT equilibria is decreasing at the level of

[5] A deterministic version of g-TFT corresponds to the well-known TFT strategy with a slight modification for the case of imperfect private monitoring, according to which, a player always mimics her opponent's action choice in the previous round by making the cooperative (defective) action choice whenever she observes the good (bad) signal.

[6] Note that TFT corresponds to the g-TFT whose retaliation intensity equals unity.

monitoring accuracy. Importantly, this decreasing property plays the central role in making use of the improvement of monitoring technology and effectively saving the welfare loss caused by the monitoring imperfection. Hence, it is quite important to examine whether this property holds even empirically.

However, the retaliation intensities observed in our experimental data contradict the predictions by the abovementioned standard equilibrium theory. Contrary to the g-TFT equilibria, our subjects tend to retaliate more in the high accuracy treatment than in the low accuracy treatment. They tend to retaliate more than the level implied by the g-TFT equilibria in the high accuracy treatment, while they tend to retaliate less in the low accuracy treatment. Hence, our subjects' behavior cannot be explained by the standard theory.

As feedback from our experimental findings to theoretical development, we demonstrate an alternative theory that is more consistent with our experimental results than the standard theory. This new theory associates equilibrium behavior with psychology and bounded rationality as follows. We permit each player to be motivated by not only pure self-interest but also *reciprocity*. We permit each player to be often *naïve* enough to select actions at random. We permit the degrees of such reciprocity and naiveté to be dependent on the level of monitoring accuracy.

By incorporating reciprocity and naïveté into the equilibrium analysis, we characterize the underlying behavioral model of preferences that makes the retaliation intensity implicit from the g-TFT equilibrium increasing at the level of monitoring accuracy, that is, the model is more consistent with our experimental results. In contrast to the standard theory, the derived behavioral model guarantees the uniqueness of g-TFT equilibrium.

The rest of this paper is organized as follows. Section 2 reviews the literature. Section 3 shows the basic model. Section 4 introduces g-TFT strategy and equilibrium. Section 5 explains the experimental design. Section 6 shows the experimental results for aggregate behavior. Section 7 explains the SFEM. Section 8 shows the experimental results for individual strategies. Section 9 demonstrates the behavioral theory. Section 10 concludes.

2. Literature Review

This study contributes to the long history of research on the repeated game literature. Equilibrium theory demonstrates folk theorems in various environments, which commonly show that a wide variety of outcomes is sustained by perfect equilibria provided the discount factor is sufficient. Fudenberg and Maskin (1986) and Fudenberg, Levine, and Maskin (1994) proved folk theorems for perfect monitoring and imperfect public monitoring, respectively. These studies utilized the self-generation nature of perfect equilibria explored by Abreu (1988) and Abreu, Pearce, and Stacchetti (1990), which, however, crucially relied on the publicity of signal observations. For the study of imperfect private monitoring, Ely and Välimäki (2002), Obara (2002), Piccione (2002) explored belief-free nature as an alternative to self-generation, which motivates a player to select both cooperative action and defective action at all times. These studies showed the folk theorem in the prisoner's dilemma game in which monitoring is private and almost perfect.[7]

Subsequently, Matsushima (2004) proved the folk theorem in the prisoner's dilemma game with imperfect private monitoring by constructing review strategy equilibria as lenient behavior with long-term punishments, in which we permit the monitoring technology to be arbitrarily inaccurate. Eventually, Sugaya (2012) proved the folk theorem with imperfect private monitoring for a very general class of infinitely repeated games, by extending self-generation to imperfect private monitoring and then combining it with belief-free nature.

Closely related to the present study is Matsushima (2013), which intensively studied g-TFT strategies in a class of prisoner's dilemma games included in our model.[8] Matsushima (2013) corresponds to the theoretical benchmark of our study, showing that the class of g-TFT strategies has great advantage in equilibrium analysis compared with TFT.

The literature of experimental studies on repeated games has examined the determinants of cooperation and tested various theoretical predictions, in order to find

[7] For a survey of up to almost perfect private monitoring, see Mailath and Samuelson (2006).
[8] For other studies on g-TFT, see Nowak and Sigmund (1992) and Takahashi (1997).

clues to resolve the multiplicity problem (for a review, see Dal Bó and Fréchette, 2014). The SFEM, which is employed in this study, is frequently used in the literature on experimental repeated games (e.g., Dal Bó and Fréchette, 2011; Fudenberg, Rand, and Dreber, 2012; Aoyagi, Bhaskar, and Fréchette, 2015; Breitmoser, 2015).

Importantly, this study includes g-TFT strategies and their variants in our SFEM list. Inclusion of such stochastic action choices is scant in the literature of experimental repeated games, with the exception of Fudenberg, Rand, and Dreber (2012). However, they include only a few g-TFT strategies, aiming only to perform robustness checks for their main claims.[9] By contrast, we rigorously include many variants of g-TFT in our SFEM list to examine our subjects' retaliating policies. We also includes various non-Markovian strategies and their stochastic variants in the list. Our experimental results then support the idea that they do not retaliate every time they observe bad signals, not because they are lenient with the observation of bad signals, but because they employ (non-trivial) stochastic strategies. This finding contrasts with Fudenberg, Rand, and Dreber (2012).

The latter part of this paper shows feedback from experiments to theory by incorporating behavioral aspects in rational behavior. This could be regarded as the first study to attempt to analyze repeated game theory with more relevance to real behavior.

3. The Model

This study investigates a repeated game played by two players, that is, players 1 and 2, using a discrete time horizon. This game has a finite round-length, but the terminating round is randomly determined, and therefore, unknown to players. The component game of this repeated game is denoted by $(S_i, u_i)_{i\in\{1,2\}}$, where S_i denotes the set of all actions for player $i\in\{1,2\}$, $s_i \in S_i$, $S \equiv S_1 \times S_2$, $s \equiv (s_1, s_2) \in S$, $u_i : S \to R$, and $u_i(s)$ denotes the payoff for player i induced by action profile $s \in S$ in the component game. We assume that each player $i's$ payoff has the following additively separable form:

$$u_i(s) = v_i(s_i) + w_i(s_j) \quad \text{for all } s \in S, \text{ where } j \neq i.$$

[9] The experimental setup in Fudenberg, Rand, and Dreber (2012) is imperfect public monitoring in which the stochastic strategies have relatively less importance than the case of imperfect private monitoring.

The impact of a player's action choice on her (or his) welfare is *additively separable* from the impact of the opponent's action choice on her welfare.

At each round, two noisy signals $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ occur after their action choices are made, where Ω_i denotes the set of possible ω_i, $\omega \equiv (\omega_1, \omega_2)$, and $\Omega \equiv \Omega_1 \times \Omega_2$. A signal profile $\omega \in \Omega$ is randomly determined according to a conditional probability function $f(\cdot \mid s): \Omega \to R_+$, where $\sum_{\omega \in \Omega} f(\omega \mid s) = 1$ for all $s \in S$, and we assume *full support* in that $f(\omega \mid s) > 0$ for all $\omega \in \Omega$ and $s \in S$.

Let $f_i(\omega_i \mid s) \equiv \sum_{\omega_j \in \Omega_j} f(\omega \mid s)$, and we assume that $f_i(\omega_i \mid s)$ is independent of s_j. We denote $f_i(\omega_i \mid s_i)$ instead of $f_i(\omega_i \mid s)$. We utilize $\omega_i \in \Omega_i$ to denote *the signal for player $i's$ action choice*. Player $i's$ action choice s_i influences the occurrence of the signal for her action choice ω_i, but does not influence the occurrence of the signal for the opponent's action choice ω_j, where $j \neq i$.

We assume that monitoring is *imperfect*. In every round $t \in \{1, 2, ...\}$, each player i cannot directly observe either the action, denoted by $s_j(t) \in S_j$, that the opponent $j \neq i$ has selected, or the realized payoff profile, denoted by $u(s(t)) = (u_1(s(t)), u_2(s(t))) \in R^2$, where we denote by $s(t) = (s_1(t), s_2(t)) \in S$ the action profile selected in round t. Instead, player i observes the signal for opponent $j's$ action choice $\omega_j(t) \in \Omega_j$, where $j \neq i$, through which player i cannot directly but indirectly and imperfectly monitors opponent $j's$ action choice $s_j(t)$.[10]

We assume that monitoring is *private*. Each player cannot know what kind of signal her opponent receives about her own action choice. Hence, each player i knows her own action choice $s_i(t)$ and the signal for the opponent's action choice $\omega_j(t)$, but does not know either the opponent's action choice $s_j(t)$ or the signal for her own action choice

[10] Our specification of monitoring structure is in contrast to previous works, such as Green and Porter (1984) and Aoyagi and Fréchette (2009). These studies have commonly assumed that the distribution of a noisy signal depends on all players' action choices, while we assume the abovementioned independence.

$\omega_i(t)$.

The assumption of imperfect private monitoring, along with the assumption of the abovementioned independence, eliminates the impact of a player's observation about the opponent's observation about her action choices. This environment enables us to focus on the impact of monitoring imperfection about the opponent's action choices.

Let $h(t) = (s(\tau), \omega(\tau))_{\tau=1}^{t}$ denote the *history up to round* t. $H = \{h(t) \mid t = 0,1,...\}$ denotes the set of possible histories, where $h(0)$ implies the null history. *The average payoff for player* i *when history* $h(t) \in H$ *up to round* t *occurs* is defined as

$$U_i(h(t)) \equiv \frac{\sum_{\tau=1}^{t} u_i(s(\tau))}{t}.$$

From additive separability, this average payoff can be rewritten as

(1) $$U_i(h(t))$$

$$= \frac{1}{t}[\left|\{\tau \in \{1,...,t\} \mid s_i(\tau) = A\}\right| v_i(A) + \left|\{\tau \in \{1,...,t\} \mid s_i(\tau) = B\}\right| v_i(B)$$

$$+ \left|\{\tau \in \{1,...,t\} \mid s_j(\tau) = A\}\right| w_i(A) + \left|\{\tau \in \{1,...,t\} \mid s_j(\tau) = B\}\right| w_i(B)].$$

This study specifies the component game as a *prisoner's dilemma with symmetry and additive separability,* as follows. For each $i \in \{1,2\}$,

$S_i = \{A, B\}$, $v_i(A) = -Y$, $v_i(B) = 0$, $w_i(A) = X + Y$, and

$w_i(B) = X + Y - Z$,

where X, Y, and Z are positive integers, and

$Z > Y > 0$.

This specification implies

$u_1(A, A) = u_2(A, A) = X$,

$u_1(B, B) = u_2(B, B) = X + Y - Z$,

$u_1(A, B) = u_2(B, A) = X - Z$,

and

$u_1(B, A) = u_2(A, B) = X + Y$.

Let us call A and B the *cooperative* action and *defective* action, respectively.

Selecting cooperative action A instead of defective action B costs Y, but gives the opponent benefit Z, which is greater than Y. Note that the payoff vector induced by the cooperative action profile (A,A), that is, (X,X), maximizes welfare $u_1(s)+u_2(s)$ with respect to $s \in S$, and is greater than the payoff vector induced by the defective action profile (B,B), that is, $(X+Y-Z, X+Y-Z)$, because $Z>Y$. Note also that the defective action profile (B,B) is a dominant strategy profile and the unique Nash equilibrium in the component game.

We further specify the monitoring structure as

$$\Omega_i = \{a,b\}, \quad f_i(a \mid A) = f_i(b \mid B) = p \text{, and } 1/2 < p < 1.$$

Let us call a and b *good* and *bad* signals, respectively. The probability index p implies the level of *monitoring accuracy*. The greater p is, the more accurately each player can monitor the opponent's action choice. Inequality $p > 1/2$ implies that the probability of a good signal occurring for a player is greater when this player selects A than when she selects B.

For each history $h(t) \in H$ up to round t, we define the *frequency of cooperative action choice* A, or, the *cooperation rate*, by

$$\rho(h(t)) \equiv \frac{\left|\{\tau \in \{1,...,t\} \mid S_1(\tau) = A\}\right| + \left|\{\tau \in \{1,...,t\} \mid S_2(\tau) = A\}\right|}{2t}.$$

From (1) and symmetry, it follows that the *sum of the average payoffs, that is, the average welfare, when history* $h(t) \in H$ *up to round* t *occurs* is given by

(2)
$$\begin{aligned}
&U_1(h(t)) + U_2(h(t)) \\
&= 2[\rho(h(t))\{v_1(A) + w_1(A)\} + \{1-\rho(h(t))\}\{v_1(B) + w_1(B)\}] \\
&= 2[\rho(h(t))X + \{1-\rho(h(t))\}(X+Y-Z)] \\
&= 2[X + \{1-\rho(h(t))\}(Y-Z)].
\end{aligned}$$

From (2), *the cooperation rate* $\rho(h(t))$ *uniquely determines welfare* $U_1(h(t)) + U_2(h(t))$.

We assume *constant random termination*, in which $\delta \in (0,1)$ denotes the probability of the repeated game continuing at the end of each round t, provided this game continues up to round $t-1$. Hence, the repeated game is terminated at the end of each round $t \geq 1$ with probability $\delta^{t-1}(1-\delta)$. The expected length of the repeated game is given by $\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t$.

Throughout our experiments, we assume that

$$\delta = 0.967 \quad (=29/30),$$

implying that the continuation probability is very high. This mimics the discount factor that is sufficiently large, to support the existence of equilibria in which players collude with each other in infinitely repeated interactions.

4. Generous Tit-For-Tat Equilibrium

Let $\alpha_i \in [0,1]$ denote a stochastic action choice for player i, implying the probability that she makes the cooperative action choice A. Player $i's$ strategy in the repeated game is defined as $\sigma_i : H \to [0,1]$. According to σ_i, she selects cooperative action A with probability $\sigma_i(h(t-1))$ in each round t, provided history $h(t-1)$ up to round $t-1$ occurs. Let Σ_i denote the set of all strategies for player i. Let $\sigma \equiv (\sigma_1, \sigma_2)$ and $\Sigma \equiv \Sigma_1 \times \Sigma_2$.

The *expected payoff for player* i *induced by* $\sigma \in \Sigma$ *when the level of monitoring accuracy is given by* $p \in (0,1)$ is defined as

$$U_i(\sigma; p) \equiv (1-\delta)E[\sum_{\tau=1}^{\infty} \delta^{\tau-1} u_i(s(\tau)) \mid \sigma, p],$$

or, equivalently,

$$U_i(\sigma; p) = \frac{E[\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta) \sum_{\tau=1}^{t} u_i(s(\tau)) \mid \sigma, p]}{\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t}$$

$$= \frac{E[\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta) t U_i(h(t)) \mid \sigma, p]}{\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t},$$

where $E[\cdot \mid \sigma, p]$ denotes the expectation operator conditional on (σ, p). *The expected* *frequency of cooperative action choice* A *, that is, the expected cooperation rate, induced by* $\sigma \in \Sigma$ *when the monitoring accuracy is given by* $p \in (0,1)$, is defined as

$$\rho(\sigma; p) \equiv \frac{E[\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t\rho(h(t)) \mid \sigma, p]}{\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t}.$$

From additive separability and (2), it follows that

$$U_1(\sigma; p) + U_2(\sigma; p) == \frac{E[\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t\{U_1(h(t)) + U_2(h(t))\} \mid \sigma, p]}{\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t}$$

$$= 2[X + \{1-\rho(\sigma; p)\}(Y-Z)].$$

Hence, *the expected cooperation rate* $\rho(\sigma; p)$ *uniquely determines the sum of the expected payoffs, that is, welfare in expectation* $U_1(\sigma; p) + U_2(\sigma; p)$. A strategy profile $\sigma \in \Sigma$ is said to be *a (subgame perfect) equilibrium* in the repeated game with monitoring accuracy $p \in (0,1)$ if

$$U_i(\sigma; p) \geq U_i(\sigma_i', \sigma_j; p) \text{ for all } i \in \{1,2\} \text{ and all } \sigma_i' \in \Sigma_i.$$ [11]

We introduce a class of strategies, namely, *g-TFT strategies,* as follows. Strategy $\sigma_i \in \Sigma_i$ is said to be *g-TFT* if $(q, r(a), r(b)) \in [0,1]^3$ exists such that

$$r(a) > 0,$$

$$\sigma_i(h(0)) = q,$$

for every $t \geq 2$ and $h(t-1) \in H$,

$$\sigma_i(h(t-1)) = r(a) \text{ if } \omega_j(t-1) = a,$$

and

$$\sigma_i(h(t-1)) = r(b) \text{ if } \omega_j(t-1) = b.$$

At round 1, player i makes cooperative action choice A with probability q. In each round $t \geq 2$, player i makes cooperative action choice A with probability $r(\omega_j)$ when she observes signal $\omega_j(t-1) = \omega_j$ for the opponent's action choice in round $t-1$.

---

[11] The full-support assumption makes the distinction between Nash equilibrium and subgame perfect Nash equilibrium redundant.

We simply write $(q, r(a), r(b))$ instead of σ_i for any g-TFT strategy.

A g-TFT strategy $(q, r(a), r(b))$ is said to be an *equilibrium* in the repeated game with accuracy $p \in (0,1)$ if the corresponding symmetric g-TFT strategy profile is an equilibrium in the repeated game with accuracy p. Let us define

(3) $$w(p) \equiv \frac{Y}{\delta(2p-1)Z}.$$

Note that

$$0 < w(p) \leq 1 \text{ if and only if } \delta \geq \frac{Y}{(2p-1)Z}.$$

The following theorem shows that a g-TFT strategy $(q, r(a), r(b))$ is equilibrium if and only if the difference in cooperation rate between the good and bad signals, that is, $r(a) - r(b)$, is equal to $w(p)$.

Theorem 1: *A g-TFT strategy* $(q, r(a), r(b))$ *is an equilibrium in the repeated game with accuracy* p *if and only if*

(4) $$\delta \geq \frac{Y}{(2p-1)Z},$$

and

(5) $$r(a) - r(b) = w(p).$$

Proof: Selecting $s_i = A$ instead of B costs player i Y in the current round, whereas in the next round, she (or he) can gain Z from the opponent's response with probability $pr(a) + (1-p)r(b)$ instead of $(1-p)r(a) + pr(b)$. Since she must be incentivized to select both actions A and B, indifference between these action choices must be a necessary and sufficient condition, that is,

$$-Y + \delta Z\{pr(a) + (1-p)r(b)\} = \delta Z\{(1-p)r(a) + pr(b)\},$$

or,

$$Y = \delta Z(2p-1)\{r(a) - r(b)\},$$

implying (5). Since $r(a) - r(b) \leq 1$, $1 \geq \frac{Y}{\delta(2p-1)Z}$, that is, inequality (4) must hold.

Q.E.D.

From Theorem 1, whenever $\delta \geq \frac{Y}{(2p-1)Z}$, by setting $r(a)-r(b)=w(p)$, we can always construct a g-TFT equilibrium. This implies an advantage of g-TFT strategies over the TFT strategy in equilibrium analyses, because the profile of TFT strategies is not an equilibrium as long as $\delta \neq \frac{Y}{(2p-1)Z}$, while g-TFT equilibrium always exists whenever $\delta \geq \frac{Y}{(2p-1)Z}$.[12]

This study regards the observed difference in the cooperation rate between the good and bad signals as the intensity with which subjects retaliate against their opponents. We call this *retaliation intensity*. Note from Theorem 1 that if subjects play a g-TFT equilibrium, then the resultant retaliation intensity is approximately equal to $w(p)$.

Importantly, the retaliation intensity implied by the g-TFT equilibria $w(p)$ is *decreasing in* p. The less accurate the monitoring technology is, the more severely players retaliate against their opponents. This decreasing property is essential for understanding how players overcome the difficulty of achieving cooperation under imperfect private monitoring.

In order to incentivize a player to make the cooperative action choice, it is necessary that her opponent makes the defective action choice when observing the bad signal more often than when observing the good signal. In other words, *the retaliation intensity must be positive*. When monitoring is inaccurate, it is difficult for her opponent to detect whether the player actually makes the cooperative action choice or the defective action choice. In this case, the enhancement in retaliation intensity is necessary to incentivize the player. Hence, *retaliation intensity must be decreasing at the level of monitoring accuracy*.

[12] Note that the g-TFT equilibrium does not depend on the assumption that ω_1 and ω_2 are independently distributed. Whether a g-TFT strategy is an equilibrium is irrelevant to whether monitoring is private or public. For instance, any g-TFT strategy that satisfies (4) and (5) is an equilibrium even if each player can observe both ω_1 and ω_2, that is, even under imperfect public monitoring.

This decreasing property plays a central role in improving welfare by utilizing noisy signals as much as possible. Since monitoring is imperfect, it is inevitable that the opponent observes the bad signal even if the player actually makes the cooperative action choice. This inevitably causes welfare loss to occur, because the opponent might retaliate against the player even she actually makes the cooperative action choice. In this case, if the monitoring technology is more accurate, the opponent can well incentivize the player by being less sensitive to whether the observed signal is good or bad, thereby safely lowering the retaliation intensity. This serves to decrease the welfare loss caused by the monitoring imperfection. Hence, it is very crucial in welfare analysis to examine whether the experimental results satisfy this decreasing property.

This study experimentally evaluates retaliation intensities. We regard g-TFT strategies as the most appropriate theoretical benchmark. G-TFT strategies are straightforward stochastic extensions of the well-known TFT strategy, extended to have simple subgame perfect equilibria in imperfect private monitoring. G-TFT strategies often appear as canonical strategies in the theoretical development of equilibrium with private monitoring. Not only is the existence of equilibria guaranteed, but also, g-TFT strategies are the only equilibrium strategies that provide the simple, analytically tractable characterization. Analytical tractability is appealing in the empirical investigation of retaliation intensities.

5. Experimental Design

We conducted four sessions of computer-based laboratory experiments[13] at the Center for Advanced Research for Finance, University of Tokyo in October 2006. We recruited 108 subjects in total from the subject pool consisting of undergraduate and graduate students in various fields. Our subjects were given monetary incentives; the points earned in the experiments were converted into Japanese Yen at a fixed rate (0.6 JPY per point). In addition, our subjects were each paid a fixed participation fee of 1,500 JPY.

As demonstrated in Section 3, to simplify the structure of the game, we adopt the

[13] The experiment was programmed and conducted with z-Tree software (Fischbacher, 2007).

prisoner's dilemma with symmetry and additive separability for our component game, in which payoff parameters are characterized as $(X, Y, Z) = (60, 10, 55)$. The payoff parameters have a structure in which the cost for cooperation Y is small so that g-TFT equilibria exist even if the monitoring technology is poor. The payoff matrix employed in the experiment is displayed in Table 1. The labels on the actions and signals are presented in neutral language (i.e., the actions are labeled "A" and "B" instead of "cooperation" and "defection," and the signals are labeled "a" and "b" instead of "good" and "bad").

[TABLE 1 HERE]

The experiments have two treatments that differ with respect to monitoring accuracy; in one treatment, monitoring accuracy is high. In this treatment, the signals the player observes and the action choices by the opponent coincide with 90% chance ($p = 0.9$), and mismatch with 10% chance. We refer to this treatment as the "high accuracy treatment." The other treatment is the case in which the monitoring technology is poorer. The signals match with the opponent's action choices with only 60% chance ($p = 0.6$), which is but still significantly larger than the chance level (50%). We refer to this treatment as the "low accuracy treatment."

All subjects in the four sessions received both treatments, but the treatment order was counterbalanced to minimize order effects. The first two sessions (52 subjects) started with three repeated games of the low accuracy treatment and then proceeded to three repeated games of the high accuracy treatments. The second two sessions (56 subjects) were conducted in the reverse order, starting with the high accuracy treatment and proceeding to the low accuracy treatment. Each treatment was preceded with a short practice repeated game, consisting of two rounds, to let the subjects understand the new treatment. Subjects were randomly paired at the start of each repeated game, and the pairs remained unchanged until the end of the repeated game. Table 2 displays the summary of the treatment order, number of subjects, and game lengths in repeated games that are determined by the continuation probability explained hereafter.

[TABLE 2 HERE]

We employed constant random termination in our experiment to mimic the infinitely repeated interactions. Each repeated game was terminated in each round with a constant probability. We let the continuation probability be $\delta = 0.967(=29/30)$.

With probability $1-\delta\,(=1/30)$, the repeated game was terminated in the current round, and subjects were re-matched with a new opponent and proceeded to the next repeated game. Indeed, our subjects were not informed in advance which round was the final round for each repeated game, otherwise we would not have been able to mimic infinitely repeated games owing to "the shadow of the future" (Dal Bó, 2005). To help our subjects understand that the termination rule is stochastic and the probability of termination in each round is 1/30, we presented 30 cells (numbered from 1 to 30) on the computer screen at the end of each round; one number was selected randomly at the end of each round, and the repeated game was terminated if the number 30 was selected by chance and the 30th cell in the computer screen turned green, otherwise all the cells numbered 1 to 29 turned green simultaneously and the repeated game continued. The screen is demonstrated in Figures A.1 and A.2 in Appendix 1.

Each subject was informed of the rule of games demonstrated above and the ways in which the game would proceed on the computer screen with the aid of printed experimental instructions. The instructions were also explained aloud by a recorded voice. Moreover, using the computer screen during the experiments, our subjects were always able to look over the structure of the game (the payoff matrix and accuracy of the signals in the treatment) and the history up to the round in the repeated game, consisting of her own actions and the signals on the opponent's actions. See Appendix 6 for the experimental instructions and the images of the computer screen, translated into English from the original Japanese[14].

6. Aggregate Level Analysis of Experimental Results

[14] Some of our subjects in pilot experiments asked the experimenters a maximum number of rounds as obviously no one is able to continue experiments without any time constraints. Accordingly, we established a sufficient maximum length of a game in the main experiments. We honestly announced it in the experimental instruction. Following to the continuation probability in Section 5, we randomly generated numbers of rounds of games. Our subjects were informed of the continuation probability, but they were not informed of the exact numbers of rounds that were randomly determined.

6.1. Overall Cooperation Rates and Round 1 Cooperation Rates

[TABLE 3 HERE]

Table 3 displays the descriptive summary of the data. The entire 108 subjects made 8,864 decisions in the high accuracy treatment, and 9,144 in the low accuracy treatment. The overall frequency of cooperative choices, that is, the cooperation rate, is 0.672 in the high accuracy treatment, and 0.355 in the low accuracy treatment. The former is statistically significantly larger than the latter ($p < 0.001$, Wilcoxon matched-pair test for individual-level means). A first look at the cooperation rates suggests that our subjects cooperate more as the monitoring technology is improved. As shown in Section 4, the cooperation rate uniquely determines the welfare of players in our payoff setting (additive separable payoff structure). A larger cooperation rate in the high accuracy treatment indicates that welfare is improved in the high accuracy treatment.

To examine our subjects' attitude for cooperation more directly, we focus on round 1 cooperation rates, which are the frequency of cooperative action choices in round 1. The round 1 cooperation rates are one of the primary measures that directly reflect subjects' motivation for cooperation under the structure of the game (i.e. payoff parameters, discount factor, and most importantly in our study, monitoring accuracy), independently from the history of the incoming signals in the latter rounds, and also independently from the behavior of the opponent matched in the repeated game.

[TABLE 4 HERE]

Table 4 presents the frequency of cooperation in round 1. The round 1 cooperation rate is 0.781 in the high accuracy treatment, and 0.438 in the low accuracy treatment, which is significantly smaller ($p < 0.001$) than the former. Our subjects tend to start repeated games with cooperative action choices in the high accuracy treatment; however, their motivation for cooperation is discouraged as the noise in the signal is increased, to the extent that they start with cooperative actions with less than 50% chance. This result is somewhat surprising given that the experimental parameters (i.e., payoff parameters and discount factor) are conductive to the g-TFT cooperative equilibria even in the low accuracy treatment. Despite the favorable environment for cooperation, our results indicate that, from the start, a non-negligible number of our subjects become reluctant to cooperate in the low accuracy treatment.

In addition, these results imply that our subjects differentiate their strategies between the two treatments. Reacting to the change of the signal quality, perhaps they switch their strategies from cooperative ones in the high accuracy treatment to less cooperative ones in the low accuracy treatment. We explore the specific strategies our subjects follow in each treatment in Section 8.

6.2. Signal-Contingent Cooperation Rates

Table 4 also displays signal-contingent cooperation rates. The frequency of cooperative actions after observing a good signal, denoted by $r(a;p)$, computed as the simple mean of entire choices, is 0.852. In addition, Table 4 reports the alternative value, which is the mean of individual-level means, concerned with the possibility that the behavior of the cooperative subjects might be over-represented in the simple mean of choices.[15] The mean of the individual-level means is 0.788, which is smaller by 0.064 than the simple mean of choices, which implies that over-representation might exist. However, both measures are consistently high, reaching around 0.8, thereby indicating that our subjects are quite cooperative when they observe a good signal in the high accuracy treatment.

In the low accuracy treatment, the cooperation rate after observing a good signal is not as high as observed in the high accuracy treatment. The simple mean of the cooperative choices is 0.437, and the mean of individual-level means is 0.423. Both low values in the cooperation rates of around 0.43 indicate that, similarly to the case of the round 1 cooperation rate, our subjects are reluctant to cooperate even after observing good signals in the low accuracy treatment, as opposed to the case in the high accuracy

[15] Since the subjects who are cooperative might observe good signals more often and adopt cooperative actions more often, their cooperative choices might be over-represented in the computation of cooperation rates contingent on a good signal.

treatment. The direct comparison of the cooperation rates between the two treatments indicates that the cooperation rate in the high accuracy treatment is larger than that in the low accuracy treatment ($p < 0.001$ for both the simple mean and the mean of individual-level means).

As for the case of the cooperation rates after observing a bad signal, denoted by $r(b; p)$, the simple mean of cooperative choices in the high accuracy treatment is 0.344. Again, we report the mean of individual-level means concerned with the over-representation of the subjects who tend to retaliate more. The value is 0.443, which is larger than the simple mean of choices by 0.104, demonstrating that the over-representation might be the case here. However, both measures are still below 50%, indicating that our subjects tend to defect rather than cooperate after observing a bad signal in the high accuracy treatment.

This tendency of our subjects to defect after observing a bad signal is more apparent in the low accuracy treatment. The simple mean of cooperative actions over the entire choices is 0.272 and the mean of the individual-level means is 0.279, which are consistently smaller than those in the high accuracy treatment ($p < 0.001$ for both means). Observing a bad signal, our subjects tend to defect more in the low accuracy treatment than in the high accuracy treatment.

The overall picture of the round 1 cooperation rates and the signal-contingent cooperation rates shown above robustly demonstrates that our subjects take more cooperative actions under the better signal quality, irrespective of signals they observe.

RESULT 1-a: Our subjects tend to cooperate more in the high accuracy treatment than in the low accuracy treatment, adapting their strategies according to the signal accuracy.

6.3. Retaliation Intensity

[TABLE 5 HERE]

Next, we focus on retaliation intensities, which are one of the primary concerns in our study. We examine whether the observed retaliation intensities coincide with the values implied by the g-TFT equilibria ($w(p)$). Table 5 presents the retaliation intensities at the aggregate level. In the high accuracy treatment, the retaliation intensity $r(a;0.9)-r(b;0.9)$ is 0.508 in the mean of entire choices and 0.352 in the mean of individual-level means. Again, the discrepancy might be due to over-representation. However, both measures of retaliation intensity consistently differ from 0 statistically significantly ($p < 0.001$ for both measures). These results indicate that our subjects use signal-contingent information in their action choices, perhaps attempting to incentivize the opponent to cooperate.

In comparison to the theoretical values implied by the g-TFT equilibria, both measures of the retaliation intensity in the high accuracy treatment are statistically significantly larger than the theoretical values ($w(0.9)=0.235$, $p < 0.001$ for both measures). Thus, empirically, our subjects tend to rely on stronger punishments that are more than enough to incentivize the opponents to take cooperative actions in the high accuracy treatment.

On the other hand, in the low accuracy treatment, the observed retaliation intensity $r(a;0.6)-r(b;0.6)$ is 0.165 in the simple mean of entire choices, and 0.144 in the mean of individual-level means. The two measures consistently differ from 0 significantly ($p < 0.001$ for both measures), which demonstrates that, similarly to the case of the high accuracy treatment, our subjects use the signal-contingent information even with the poorer monitoring technology. However, unlike the case of the high accuracy treatment, the observed retaliation intensity in the low accuracy treatment is smaller than the level implied by the g-TFT equilibria. Both measures of retaliation intensity are outstandingly smaller than the theoretically implied value $w(0.6)=0.94$ (both for $p < 0.001$). Although our subjects do retaliate according to the signals in the low accuracy treatment, the strength of the retaliation is far below the level to incentivize opponents to cooperate, allowing the opponents to defect permanently to pursue larger payoffs.[16]

[16] There might be concern that the seemingly weaker retaliation intensities observed in this study do not necessarily imply weak retaliation policies of our subjects, since our subjects might use long-term (multi-round) punishments; long-term punishing strategies punish opponents even observing a good signal during punishing phases, which could lower the retaliation intensity computed here. However, our analysis in Section 8 indicates that only a marginal number of our subjects could adopt long-term punishing strategies, and hence, the effect of long-term punishments is minimal.

Not only are the retaliation intensities inconsistent with the values implied by the g-TFT equilibria, but also, the deviation is systematic. A direct comparison of the retaliation intensities between the two treatments indicates that the retaliation intensity in the high accuracy treatment is larger than that in the low accuracy treatment ($p < 0.001$). This is diametrically opposite to the implication by the g-TFT equilibria in which people should employ weaker retaliating policies in the high accuracy treatment than in the low accuracy treatment.

RESULT 2-a: Our subjects tend to retaliate more than the level implied by the g-TFT equilibria in the high accuracy treatment, while they tend to retaliate less in the low accuracy treatments. Moreover, contrary to the implications of the canonical theory, the retaliation intensity is larger with the improved monitoring technology.

One might concern that, as the signal becomes less accurate, some subjects simply regard the signal as being uninformative, and hence neglect the signal. Those subjects would have higher retaliation intensities in the high accuracy treatment than in the low accuracy treatment.

If this were the case, the whole subjects would divide into two groups; one group of subjects would have zero retaliation intensities in the low accuracy treatment (i.e., higher retaliation intensities in the high accuracy treatment than in the low accuracy treatment), while the other group of subjects would have sufficiently high retaliation intensities in the low accuracy treatment.

Our data do *not* support this hypothetical case. Indeed, only a small number of our subjects have higher retaliation intensities in the low accuracy treatment than in the high accuracy treatment. Many subjects have retaliation intensities that are non-negligible but insignificant in the low accuracy treatment. A detailed discussion is provided in Section 8, in which we investigate individual strategies.

6.4. Reliance on Long Memories

In relation to the behavioral difference of our subjects along with signal qualities, it is interesting to examine whether our subjects tend to rely on longer histories of signals (i.e., signals two periods ago). Given the theories in review strategies (Radner, 1986; Matsushima, 2004; Sugaya, 2012), people might rely on signals in longer histories to compensate for the informational disadvantages of poorer monitoring technologies.

To test whether our subjects rely on information in a signal occurring two periods ago, here, we fit the data with a probabilistic linear regression model, regressing the action choices (a dummy variable that takes 1 if our subjects play cooperation) on all memory-1 histories, which consist of a signal and an action in the previous round, and further include information of a signal two periods ago into the set of explanatory variables to the extent that non-singularity holds without intercept.[17] The regression coefficients on the signal two periods ago capture the additional impact on cooperation probabilities from the signal. The standard error is computed by a cluster-bootstrap (subject and repeated game level) to control subject heterogeneity, which is also used to compute the p-value.

[TABLE 6 HERE]

Table 6 displays the regression results. Contrary to the abovementioned expectations, the regression coefficients on the information two periods ago are exclusively significant only in the high accuracy treatment, and none is significant in the low accuracy treatment. Even in the high accuracy treatment, the size of the coefficients is only marginal, as the maximum impact on cooperative choices is only 13.2%.

RESULT 3-a: There is no evidence that our subjects tend to review longer periods in the low accuracy treatment. Although their actions partly depend on the information two periods ago in the high accuracy treatment, the dependencies are marginal at the aggregate level.

6.5. Impact of Experiences

[17] We borrow this approach from Breitmoser (2015).

Several existing studies have reported that the frequency of cooperation changes as people experience playing repeated games (see discussion in Dal Bó and Fréchette, 2014). Since welfare is uniquely determined by the overall cooperation rates in our payoff setting (the additive separability payoff structure), a shift of overall cooperation rates implies a shift of welfare of the two players.

To examine the impact of the experience of repeated games on overall cooperation rates, we perform a reduced-form, linear regression analysis, which is explained in detail in Appendix 2. The results indicate that, although there are some experience effects on action choices, the sizes of the effects are not remarkably large in our data. In addition, we find qualitatively similar results on signal-contingent cooperation rates, which are also reported in Appendix 2.

Furthermore, we investigate the effect of experience on retaliation intensities. Here, we perform a similar, reduced-form regression analysis, explained in Appendix 2 in detail. The results indicate that the retaliation intensities do not change remarkably as our subjects gain experience.

7. Estimation of Individual Strategies—Methodology

In this and the following Section 8, we present the direct estimation of individual strategies of our subjects. Given the recent consensus in the literature of experimental repeated game that there is substantial heterogeneity of strategies employed by subjects (for a review, see Dal Bó and Fréchette, 2014), we do not specifically pick only a single class of strategy (i.e., g-TFT) and fit the data to determine the model parameters. Rather, we list various strategies our subjects could adopt, and estimate the frequency with which each strategy emerges among our subjects to assess the prevalence of the strategies. The primary goal of the exercise here is to verify our findings in Section 6 and to perform more detailed analyses on g-TFT strategies and retaliation intensities from the viewpoint of individual strategies.

The methodology we employ here is the SFEM developed by Dal Bó and Fréchette (2011). The SFEM is a maximum likelihood estimation (MLE) of a finite mixture model of strategies that subjects individually use, in which model parameters to be estimated are the frequencies of each strategy emerging among subjects, and parameter γ which controls stochastic mistakes of action choices or implementation error, whose probability of occurrence is $1/(1+\exp(1/\gamma))$. The details of the computation of the likelihood are provided in Appendix 3. The underlying assumption is that each subject keeps employing a specific strategy across all repeated games in each treatment. The validity of this method is verified in the Monte Carlo simulations in Fudenberg, Rand, and Dreber (2012), in which, closely related to our study, the authors correctly dissociate g-TFT players from players employing other strategies.

In Appendix 4, we perform robustness checks of the SFEM estimates using only the final two repeated games. The SFEM estimates have few changes in each treatment, which would not occur if our subjects systematically changed their strategies across repeated games. This is consistent with our findings in Section 6, where we do not find remarkable changes in either the cooperation rates or the retaliation intensities as our subjects gain experience across repeated games.

[TABLE 7 HERE]

In the SFEM framework, the strategy set considered in the estimation is pre-specified. Given the difficulty of covering all possible sets of strategies, we include only the strategies that share significant proportions in existing studies on experimental repeated games as well as g-TFT, which is our primary focus. Table 7 displays the list of the strategies in our SFEM. The list includes TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim (trigger strategy),[18] Grim-2, Grim-3, always cooperate (All-C), and always defect (All-D), which are typically listed in the literature of infinitely repeated games in imperfect monitoring (e.g., Fudenberg, Rand, and Dreber, 2012).[19] Among them, All-D is a non-cooperative strategy, while other strategies (TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim,

---

[18] Here, the definition of Grim strategy is modified to cover the private monitoring case in which no common signals are observable. The player starts to keep playing defection if she observed a bad signal or played defection in the previous round. Note that she could mistakenly play defection before the "trigger" is pulled, since the implementation error of action choices are allowed in the SFEM framework.

[19] The literature has often added D-TFT to the strategy set, in which defection is played in round 1, followed by TFT from Round 2. However, we find no significant frequency of D-TFT in either treatment in our SFEM estimates, even if we include D-TFT.

Grim-2, Grim-3, and All-C) are regarded as cooperative strategies that play cooperation at the start of each repeated game, and keep cooperating unless they believe the opponent might switch to defect. TF2T, TF3T, 2TF2T, Grim-2, and Grim-3 are "lenient" strategies (Fudenberg, Rand, and Dreber, 2012) that start punishing only after observing several consecutive occurrences of bad signals. TF2T (TF3T) retaliates once after observing two (three) consecutive bad signals, and 2TF2T retaliate twice after observing two consecutive bad signals, corresponding to a simple form of so-called "review strategies" (lenient strategies with long-term punishments in the proof of the limit folk theorem (see Matsushima, 2004; Sugaya, 2012). Grim-2 (Grim-3) is a lenient variant of Grim strategy, which triggers continuous defection after observing two (three) consecutive deviations from (a, A), the combination of good signal from the opponent and own cooperative choice.

Importantly, motivated by the theoretical importance of g-TFT in imperfect private monitoring, we include g-TFT in the list. Inclusion of g-TFT is scant in the literature of experimental repeated games. The exception is Fudenberg, Rand, and Dreber (2012), who found that at least a certain share of their subjects follow g-TFT even in imperfect public monitoring in which g-TFT is relatively less important. However, the authors' discussion on g-TFT is incomplete since they included g-TFT only with the aim of performing robustness checks for their main claims, which are less relevant to g-TFT, and thus, they left further discussions unanswered. Rather in this study, to address the implications of g-TFT and associated retaliation intensities more rigorously in imperfect private monitoring, we add many variants of g-TFT in our SFEM strategy to dissociate strategies with various retaliation intensities.

Following Fudenberg, Rand, and Dreber (2012), we pre-specify probabilities of retaliation, or equivalently, probabilities of cooperation after observing a bad signal in g-TFT (i.e., $r(b)$) in our strategy set. We allow the probabilities of cooperation given a bad signal to take nine distinct values in increments of 12.5%, that is, 100%, 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5%, and 0%. Moreover, to cover the case in which our subjects might play defections even after observing a good signal (i.e., $r(a) < 1$), we allow stochastic defections given a good signal. The probabilities of cooperation given a good signal (i.e. $r(a)$) are allowed to take the nine distinct values in increments of 12.5%, that is, 100%, 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5%, and 0%. Here, g-TFT- $r(a)$ - $r(b)$ denotes the g-TFT that plays cooperation stochastically after observing a good signal with probability $r(a)$, and adopts cooperation stochastically after observing a bad signal with probability $r(b)$.[20] We list all possible combinations of $r(a)$ and $r(b)$ in g-TFT in our strategy set as long as the g-TFT has a non-negative retaliation intensity (i.e., $r(a) \geq r(b)$). Specifically, we refer to the g-TFT strategies playing cooperation with constant probabilities r irrespective of signals as random strategies (equivalent to g-TFT- r - r), denoted by Random- r , as primitive, signal non-contingent, zero retaliation variants of g-TFT, including All-C and All-D as special cases. We regard a g-TFT strategy as non-cooperative if the g-TFT is less cooperative than Random-0.5. Specifically, a g-TFT is regarded as non-cooperative if $r(a)$ and $r(b)$ in the g-TFT are no more than 0.5; otherwise, the g-TFT strategies are cooperative.

In addition to including these g-TFT strategies in the list, we add a family of g-2TFT as strategies that perform even stronger punishments than TFT. The motivation comes from our earlier analysis in Section 6, in which we find that our subjects aggregately adopt stronger retaliation intensities than the level implied by the standard theory in the high accuracy treatment. The family of g-2TFT strategies (g-2TFT- r) allows the second retaliations to be stochastic (play cooperation with probability r in the second punishment) as the generous variants of 2TFT.[21] Conservatively, we also include a family of g-TF2T (g-TF2T- r) as the generous variants of TF2T, which allow stochastic punishments if two consecutive bad signals occur (play cooperation with probability r after observing two consecutive bad signals).[22]

[20] For simplicity, we assume the probability of playing cooperation in round 1 coincides with the choice probability given a good signal (i.e., $r(a)$).

[21] Unlike g-TFT, in g-2TFT, we do not allow defections until the punishing phases start. Allowing defections outside of punishing phases starts reducing the retaliation intensities, which is contrary to the motivation for employing stronger (multi-round) punishments than the punishment attainable in only one round. In addition, the strategies are assumed to play cooperation in round 1 similarly to TFT, as only multi-round punishing variants of TFT.

[22] Unlike g-TFT, in g-TF2T, we do not allow defections until the punishing phases start, since defections outside of punishing phases are contrary to the motivation for employing lenient strategies that allow "giving the benefit of the doubt" to an opponent on the first defection (Fudenberg, Rand, and Druber, 2012). For the same reason, the strategies are assumed to play cooperation in round 1.

8. Estimation of Individual Strategies—Results

8.1. Cooperative Strategies and Non-Cooperative Strategies

[Table 8 HERE]

[Table 9 HERE]

Table 8 presents the estimates for the frequencies of strategies our subjects follow, and Table 9 displays the aggregated frequencies. First, we focus on the share of cooperative strategies in both treatments. Our findings in Section 6 suggest that our subjects adopt cooperative strategies in the high accuracy treatment, but do not necessarily adopt cooperative strategies in the low accuracy treatment. Here, we further examine the specific shares of the cooperative strategies.

In the high accuracy treatment, cooperative strategies (i.e., strategies other than All-D, Random with a cooperation rate of no more than 0.5, and g-TFT that is less cooperative than Random-0.5 (g-TFT-0.5-$r(b)$, g-TFT-0.375-$r(b)$, g-TFT-0.25-$r(b)$, and g-TFT-0.125-$r(b)$))exceed non-cooperative strategies (i.e., All-D, Random with a cooperation rate of no more than 0.5, and g-TFT that is less cooperative than Random-0.5 (g-TFT-0.5-$r(b)$, g-TFT-0.375-$r(b)$, g-TFT-0.25-$r(b)$, and g-TFT-0.125-$r(b)$)). The share of cooperative strategies is 83.9%, and that of the non-cooperative strategies is the remainder, 16.1%. The latter share is statistically significantly smaller than the former ($p < 0.001$). Although there are considerable heterogeneities in the strategies our subjects follow, as Table 8 shows, most of our subjects adopt cooperative strategies in the high accuracy treatment.

On the contrary, our subjects tend to adopt non-cooperative strategies more in the low accuracy treatment. The share of cooperative strategies drops to 34.7% and the non-cooperative strategies' share is 65.3%. The share of non-cooperative strategies in the low accuracy treatment is statistically significantly larger than that in the high accuracy treatment ($p < 0.001$). Individually, the share of All-D is the most, that is, 19%.

The finding that the share of non-cooperative strategies is considerable in the low accuracy treatment is consistent with the previous finding in Section 6 that the round 1 cooperation rate remains low, reaching only 43.3% in the low accuracy treatment. Here, our SFEM estimates further indicate that indeed more than half of our subjects follow non-cooperative strategies in the low accuracy treatment.

This result regarding the large share of non-cooperative strategies is somewhat surprising given that the payoff parameters and discount factor are conductive to the g-TFT cooperative equilibria even in the low accuracy treatment. The complexity of the equilibrium cooperative strategy is not the primary impediment to follow cooperative strategies, since the equilibrium cooperative strategy in the low accuracy treatment is approximately TFT in our experimental setup, which is quite simple to implement. Rather, our result implies that the poor signal quality in the low accuracy treatment somewhat strongly discourages our subjects from adopting cooperative strategies.

With respect to this point, Fudenberg, Rand, and Dreber (2012) reported in their imperfect public monitoring experiment that the frequencies of cooperative strategies drop significantly as the level of noise is increased from 1/16 to 1/8. Our study demonstrates that, even by drastically changing signal noise from 1/10 to 4/10 in imperfect private monitoring, similarly, the poorer signal quality indeed discourages our subjects from adopting cooperative strategies even if the experimental parameters are conductive for cooperation.

RESULT 1-b: Most of our subjects adopt cooperative strategies in the high accuracy treatment, while in the low accuracy treatment, many of our subjects adopt non-cooperative strategies.

8.2. Proportion of G-TFT Strategies

Second, motivated by the theoretical arguments in Section 4, we focus on the share of g-TFT. Our SEFM estimates in Tables 8 and 9 indicate that there is a substantial proportion of the family of g-TFT in our imperfect private monitoring. In the high accuracy treatment, g-TFT-1-0.5 individually shares the most over the entire set of strategies, namely, 17.1%, followed by another g-TFT strategy, namely, g-TFT-0.875-

0.25 (10.7%). The total share of the g-TFT family (g-TFT-$r(a)$-$r(b)$), including TFT, but excluding signal non-contingent variants of g-TFT (i.e. All-C (g-TFT-1-1), All-D (g-TFT-0-0), and Random-r (g-TFT-r-r)), is as large as 70.6%. The extended family of g-TFT, which includes the signal non-contingent, primitive variants of g-TFT, comprises 76.8% . Moreover, the further extended family of g-TFT, which includes the family of g-2TFT as multi-round punishing variants, comprises 77.7%. As shown by these large numbers, most of our subjects follow one of the strategies in the class of g-TFT.

The share of g-TFT is also substantially large in the low accuracy treatment. The share of the g-TFT family is 55.6%, while the extended family with signal non-contingent variants comprises 94.6%, and the further extended family with multi-round punishing variants comprises 96.2%. Regardless of the treatment, we find a substantial share of our subjects following a strategy in the family of g-TFT.

This finding indicates that their decisions on retaliation largely depend on a single occurrence of a bad signal (c.f., lenient strategies). This finding is consistent with our previous finding in Section 6 that the action choices only marginally depend on the signals occurring two periods ago in both treatments.[23]

RESULT 4: Our SFEM estimates indicate that the family of g-TFT comprises a substantial proportion among the strategies our subjects follow in both treatments.

8.3. Retaliation Intensity

In this subsection, observing that many of our subjects follow g-TFT, we turn to the issue of retaliation intensities, one of the primary focuses of our study. Here, we address what proportion of our subjects adopts retaliation intensities that are consistent with g-

[23] The result that our SFEM estimates find many of our subjects playing g-TFT in our imperfect private monitoring is seemingly less consistent with the finding of Fudenberg, Rand, and Dreber (2012), which found only a small proportion of their subjects playing g-TFT in their imperfect public monitoring. However, we are not able to address the exact factors behind the discrepancy between our results and theirs, since our experimental settings are different in many aspects, such as payoff parameters, discount factor, and signal accuracy. Perhaps most importantly, our setting is a private monitoring environment in which g-TFT plays an important role as a benchmark equilibrium strategy. Nonetheless, as discussed later in Section 8.5, similarly to Fudenberg, Rand, and Dreber (2012), we find a certain proportion of our subjects following lenient strategies in the high accuracy treatment.

TFT equilibria. Recall that the level of retaliation intensity implied by the g-TFT equilibria is 0.235 in the high accuracy treatment. Approximately, g-TFT-1-0.75, g-TFT-0.875-0.625, g-TFT-0.75-0.50, g-TFT-0.625-0.37.5, g-TFT-0.50-0.25, g-TFT-0.375-0.12.5, and g-TFT-0.25-0 have this intensity of retaliation in our strategy set. However, our SFEM estimates in Tables 8 and 9 indicate that the joint share of these strategies is only 5.6% which does not significantly differ from 0 (p = 0.319), implying that very few of our subjects follow equilibrium g-TFT strategies in the high accuracy treatment.

This finding also holds in the low accuracy treatment. In the low accuracy treatment, the retaliation intensity implied by the g-TFT equilibria is much larger than that in the high accuracy treatment, which is 0.94. The retaliation intensity in TFT is approximately equivalent to the level implied by the g-TFT equilibria. However, the share of TFT in the low accuracy treatment is only 2.7%, which is not significantly different from 0 (p = 0.221). These results demonstrate that, even though many of our subjects follow one of the g-TFT strategies, almost none follow the retaliation intensities implied in the g-TFT equilibria in both treatments.

Observing that almost none of our subjects follow the retaliation intensities implied by the g-TFT equilibria, we further address how they tend to deviate from the theory. In Section 6, we find that the aggregate level of retaliation intensity in the high accuracy treatment is larger than that implied by the g-TFT equilibria. Next, we examine whether similar, consistent results emerge here in the SFEM estimates.

Our SFEM estimates in Tables 8 and 9 indicate that the group of stronger retaliation variants of g-TFT (g-TFT strategies with retaliation intensities more than 0.25, i.e., g-TFT-1-0.625/0.5/0.375/0.25/0.125/0, g-TFT-0.875-0.5/0.375/0.25/0.125/0, g-TFT-0.75-0.375/0.25/0.125/0, g-TFT-0.625-0.25/0.125/0, g-TFT-0.5-0.125/0, g-TFT-0.375-0, and g-2TFT-0.875/0.75/0.625/0.5/0.375/0.25/0.125/0) jointly comprises 54.4% in the high accuracy treatment. Previously, we find slightly more than 70% of our subjects follow g-TFT. However, the results here imply that roughly three-fourths of them retaliate more strongly than the g-TFT equilibria require. Indeed, the share of the weaker retaliation variants of g-TFT, even including signal non-contingent strategies, merely reaches 17.7% (g-TFT-1-0.875, g-TFT-0.875-0.75, g-TFT-0.75-0.625, g-TFT-0.625-0.5, g-TFT-0.5-0.375, g-TFT-0.375-0.25, g-TFT-0.25-0.125, g-TFT-0.125-0, and signal non-contingent, zero retaliation strategies, which are All-C, All-D, and Random- r), which is statistically significantly smaller than the share by the group of stronger retaliation variants ($p < 0.001$).

On the other hand, in the low accuracy treatment, we previously find weaker retaliation intensities at the aggregate level in Section 6. Now, we examine how many of our subjects follow the weaker retaliation strategies in our SFEM estimates. Our SFEM estimates indicate that the group of weaker retaliation variants of g-TFT (g-TFT other than TFT and signal non-contingent, zero retaliation variants of g-TFT, which are All-C, All-D, and Random- r) jointly comprise 93.5%. Again, we find that 96.2% of our subjects adopt strategies in g-TFT, however, they are predominantly weaker retaliation variants of g-TFT. On the contrary, strong retaliation variants of g-TFT (here, the multi-round punishing variants of TFT are g-2TFT-0.875/0.75/0.625/0.5/0.375/0.25/0.125/0) comprise less than 0.1%, which is statistically significantly smaller than the share by the weaker retaliation variants ($p < 0.001$).

Previously, in Section 6, we find that the mean retaliation intensities deviate from the values implied by the g-TFT equilibria at the aggregate level systematically. Here, we further examine whether the finding still holds even if we restrict our attention to the behavior of g-TFT players rather than the aggregate behavior of all the strategies.

To focus on the behavior of the g-TFT players, we compute the mean retaliation intensities conditional on the g-TFT strategies (including All-C, All-D, and Random- r). The conditional mean retaliation intensity in the high accuracy treatment is 0.426 (s.d. 0.033), which is significantly larger than the value predicted by the g-TFT equilibria (0.235, $p < 0.001$). In the low accuracy treatment, the mean retaliation intensity is 0.148 (s.d. 0.025), which is significantly smaller than the value implied by the g-TFT equilibria (0.94, $p < 0.001$). A direct comparison of the two mean retaliation intensities demonstrates that the conditional mean of retaliation intensities in the high accuracy treatment is significantly larger than that in the low accuracy treatment ($p < 0.001$). Even restricted to g-TFT players, their behavior systematically deviates from the theoretical predictions, similarly to the findings in Section 6.

RESULT 2-b: Our SFEM estimates indicate that, in both treatments, only a small number of our subjects follow the retaliation intensities implied by the g-TFT equilibria. In the high accuracy treatment, the share of stronger retaliation variants of g-TFT outweighs that of weaker retaliation variants. In the low accuracy treatment, the share of weaker retaliation variants of g-TFT outweighs that of stronger retaliation variants. Moreover, the mean retaliation intensity among g-TFT players is larger in the high accuracy treatment than the value implied by the g-TFT equilibria, and is smaller in the low accuracy treatment than that by the g-TFT equilibria. Furthermore, the mean retaliation intensity by the g-TFT strategies is larger in the high accuracy treatment than in the low accuracy treatment, which is contrary to the theoretical implications by the g-TFT equilibria.

As briefly discussed in Section 6, one might concern that some subjects simply regard the signal as being uninformative and neglect the signal as the signal accuracy becomes deteriorated. If this occurred, the whole subjects would divide into two groups; one group would have zero retaliation intensities in the low accuracy treatment, while the other group would have sufficiently high retaliation intensities in the low accuracy treatment.

In order to clarify whether our experimental results imply the abovementioned idea, we first compute the fraction of our subjects who follow signal non-contingent, zero retaliation strategies (i.e., ALL-C, ALL-D, and Random strategies) in the low accuracy treatment from our SFEM estimates. For each subjects, we compute the Bayesian posterior for employing the signal non-contingent strategies and picked the subjects whose posterior is more than 0.95.[24] The estimates indicate that only a small number of subjects, specifically twenty-five subjects out of 108 subjects, follow the signal non-contingent, zero retaliation strategies. For the other 83 subjects, the mean of retaliation intensities (individual-level) is 0.1883, which is still far below the value implied by the g-TFT equilibria. Indeed, only a few subjects in the subsample have strong retaliation intensities in the low accuracy treatment. Those who have strong retaliation intensities more than 0.9 (i.e., the value close to that implied by the g-TFT equilibria) are only three subjects, and those who have retaliation intensities more than 0.5 are only six subjects.

Second we compute the mean of difference of retaliation intensities (individual-level) in this 83 subsample. This value is 0.1900, which is close to the original value

[24] Other threshold values such as 0.90 or 0.80 do not alter our conclusion significantly.

presented in Table 5, implying that our finding is robust for the existence of the signal non-contingent strategies.

8.4. Long-Term Punishment

Next, we focus on long-term punishing strategies. In Section 6, we observe that the aggregate level of retaliation intensity is smaller than the level implied by the g-TFT equilibria in the low accuracy treatment. Here, we address the concern that seemingly weak retaliation intensity might spuriously arise when some subjects employ long-term punishing strategies.

Our SEFM estimates in Tables 8 and 9 indicate this is not the case in our data. The share of strategies with long-term punishments (a family of 2TFT, 2TF2T, Grim, Grim2, and Grim 3) in the low accuracy treatment is less than 3%, which is statistically insignificant ($p = 0.435$), and hence, the effect of strategies with long-term punishment is minimum. This is also true in the high accuracy case, in which the joint share by the long-term punishing strategies is 6.6%, which is statistically insignificant ($p = 0.203$).

RESULT 2-c: The share of strategies with long-term punishments that could spuriously reduce the observed retaliation intensities is small in both treatments.

8.5. Lenience

Finally, we focus on the share of lenient strategies. As seen in Section 6, our subjects do not have a tendency to rely on longer histories in the low accuracy treatment. This finding suggests that the share of lenient strategies might not be larger in the poorer monitoring technologies in our imperfect private monitoring, contrary to the speculations in the field of review strategies (Radner, 1986; Matsushima, 2004; Sugaya, 2012). Here, we examine whether the share of lenient strategies becomes larger as the monitoring technology becomes poorer.

Our SFEM estimates in Tables 8 and 9 indicate that, in the low accuracy treatment, none of the individual shares by the lenient (review) strategies (the family of g-TF2T, including TF2T, TF3T, 2TF2T, Grim-2, and Grim-3) significantly differ from 0. Even jointly, they comprise only 3.8%, which is not significantly different from 0 (p = 0.359). The lenient strategies do not share a remarkable proportion in the low accuracy treatment. On the contrary, in the high accuracy treatment, the share of lenient strategies rises to 21.4% which is significantly different from 0 (p = 0.023) and is marginally significantly larger than that in the low accuracy treatment (p = 0.055). Thus, we conclude that there is no tendency to adopt more lenient strategies in the poorer monitoring technologies.

RESULTS 3-b: Our SFEM estimates provide no evidence for a larger share of lenient strategies in the low accuracy treatment. Rather, the share is negligible only in the low accuracy treatment, while it is approximately 20% in the high accuracy treatment.

9. Feedback to Theory

Our experimental results indicate that the cooperation rate is greater in the high accuracy treatment, a substantial proportion of the subjects play g-TFT strategies, and the retaliation intensity is greater in the high accuracy treatment.

As a feedback from these experimental findings to the theoretical development, we demonstrate an alternative theory, which is more consistent with observed behavior than the standard theory. This section ignores the heterogeneity of strategies the subjects employed. We replace this heterogeneity with the common knowledge assumption about the strategy that players employ.

The purpose of this section is to associate our subjects' behavior with their aspects of psychology and bounded rationality. More precisely, we permit each player to be motivated by not only pure self-interest but also *reciprocity*. In addition, we permit each player to be often *naïve* enough to select actions at random. We further permit the degrees of such reciprocity and naiveté to be dependent on the level of monitoring accuracy.

By incorporating reciprocity and naïveté into equilibrium analysis, we characterize the *underlying behavioral model of preferences* that makes the retaliation intensity implied by the g-TFT equilibria increasing at the level of monitoring accuracy, that is, more consistent with our experimental findings.

In order to focus on the incentives in the second and later rounds, this section simply writes $(r(a), r(b))$ instead of $(q, r(a), r(b))$ for any g-TFT strategy. We fix an arbitrary $\underline{p} \in (\frac{1}{2}, 1)$ as the lower bound of monitoring accuracy.

9.1. Behavioral Model

Consider an arbitrary accuracy-contingent g-TFT strategy, which is denoted by

$$(r(a;p), r(b;p))_{p\in(\underline{p},1)}.$$

For each level of monitoring accuracy $p \in (\underline{p}, 1)$, a player makes stochastic action choices according to the g-TFT strategy $(r(a), r(b)) = (r(a;p), r(b;p))$. We assume that the player selects both actions A and B with positive probabilities, that is,

(6) $$0 < r(a;p) < 1 \text{ and } 0 < r(b;p) < 1.$$

We introduce *naïveté* as follows. In every round, with probability $2\varepsilon(p) \in [0, \frac{1}{2}]$, the player naively or randomly selects between actions A and B, that is,

$$\min[r(a;p), 1 - r(a;p), r(b;p), 1 - r(b;p)] \geq \varepsilon(p).$$

With the remaining probability $1 - 2\varepsilon(p)$, the player makes the action choice in a more conscious manner.[25] [26]

We introduce *reciprocity* as follows. Suppose that the player observes signal a, that is, the good signal for his opponent. He feels guilty when he selects the defective action despite the observation of the good signal. In this case, he can save the psychological cost $c(a;p) \geq 0$ by selecting the cooperative action. Hence, the

[25] In order to calm the tense relationship between rationality and empirical data, economic theory and empirics have used stochastic choice models, such as logit and probit models that incorporate random error into the equilibrium analysis. For instance, in the model of quantal response equilibrium, it is assumed that the deviation errors from the optimal action choice are negatively correlated with the resultant payoffs. See Goerree, Holt, and Palfrey (2008) for a review. By contrast, this section assumes that the deviation errors induced by naïveté are independent of either the resultant payoff or the observed signal but depend on the level of monitoring accuracy.

[26] The general experimental literature has often pointed out that social preference facilitates cooperation. See Güth, Schmittberger, and Schwarze (1982), Berg, Dickhaut, and McCabe (1995), and Fehr and Gächter (2000). The literature assumes that preferences depend on various contexts of the game being played. See Rabin (1993), Charness and Rabin (2002), Falk, Fehr, and Fischbacher (2003), Dufwenberg and Kirchsteiger (2004), and Falk and Fishbacher (2005). This study makes the relevant context parameterized by the level of monitoring accuracy.

instantaneous gain from selecting action B is equal to $Y - c(a;p)$, while the resultant future loss is equal to $\delta Z(2p-1)\{r(a;p)-r(b;p)\}$. From (6), we require the following properties as one part of the equilibrium constraints in behavioral theory:

(7) $$[Y - c(a;p) > \delta Z(2p-1)\{r(a;p)-r(b;p)\}] \Rightarrow [r(a;p) = \varepsilon(p)],$$

and

(8) $$[Y - c(a;p) < \delta Z(2p-1)\{r(a;p)-r(b;p)\}] \Rightarrow [r(a;p) = 1-\varepsilon(p)].$$ [27]

Next, suppose that the player observes signal b, that is, the bad signal for his opponent. He is ill-tempered when he selects the cooperative action despite the observation of the bad signal. In this case, he can save the psychological cost $c(b;p) \geq 0$ by selecting the defective action. Hence, the instantaneous gain from selecting action B is equal to $Y + c(b;p)$, while the resultant future loss is equal to $\delta Z(2p-1)\{r(a;p)-r(b;p)\}$. From (6), we require the following properties as the other part of this section's equilibrium constraints:

(9) $$[Y + c(b;p) > \delta Z(2p-1)\{r(a;p)-r(b;p)\}] \Rightarrow [r(b;p) = \varepsilon(p)],$$

and

(10) $$[Y + c(b;p) < \delta Z(2p-1)\{r(a;p)-r(b;p)\}] \Rightarrow [r(b;p) = 1-\varepsilon(p)].$$

We define a *behavioral model* as

$$(\varepsilon(p), c(a,p), c(b;p))_{p\in(\underline{p},1)}.$$

An accuracy-contingent g-TFT strategy $(r(a;p), r(b;p))_{p\in(\underline{p},1)}$ is an *equilibrium* in the behavioral model $(\varepsilon(p), c(a,p), c(b;p))_{p\in(\underline{p},1)}$ if the properties (7), (8), (9), and (10) hold for all $p \in (\underline{p},1)$.[28]

9.2. Characterization

[27] Without any substantial change, we can interpret $c(a;p)$ as the psychological benefit, in which the player feels better by selecting the cooperative action instead of the defective action after observing the good signal. The same interpretation applies to $c(b;p)$.

[28] For simplicity of argument, we eliminate the incentive constraints for the first-round action choices, and then focus on the action choices in the second and later rounds.

A player with behavioral model $(\varepsilon(p), c(a,p), c(b;p))_{p\in(\underline{p},1)}$ is *positively reciprocal* in accuracy p if $c(a;p)>0$. She is said to be *negatively reciprocal* in p if $c(b;p)>0$. She is said to be *null-reciprocal* in p if $c(a;p)=0$ and $c(b;p)=0$.

We assume that $c(a;p)$ and $c(b;p)$ are continuous in accuracy p, and

either $c(a;p)=0$ or $c(b;p)=0$.

With this assumption, we can divide players into three categories: positively reciprocal players, negatively reciprocal players, and null-reciprocal players.

According to our experimental findings, we assume that

(i) both $r(a;p)$ and $r(b;p)$ are increasing and continuous in p,

and

(ii) the retaliation intensity $r(a;p)-r(b;p)$ is increasing in p.

Since $w(p)$ is decreasing in p, it follows from (ii) that there is a critical level $\hat{p}\in[\underline{p},1]$ such that

$$r(a;p)-r(b;p)>w(p) \qquad \text{if } p>\hat{p},$$

and

$$r(a;p)-r(b;p)<w(p) \qquad \text{if } p<\hat{p}.$$

The retaliation intensity is greater than the retaliation intensity implied by the g-TFT equilibria in standard theory if the level of monitoring accuracy is better than the critical level $\hat{p}$, while the retaliation intensity is less than the retaliation intensity implied by the g-TFT equilibria in standard theory if the level of monitoring accuracy is worse than the critical level $\hat{p}$.

The following theorem shows that the abovementioned equilibrium constraints, that is, (7), (8), (9), and (10), uniquely determine the underlying behavioral model.

Theorem 2: *The accuracy-contingent g-TFT strategy* $(r(a;p), r(b;p))_{p\in(\underline{p},1)}$ *is an equilibrium in the behavioral model* $(\varepsilon(p), c(a,p), c(b;p))_{p\in(\underline{p},1)}$ *if and only if*

$$\text{(11)} \qquad r(a;p)=1-\varepsilon(p) \quad \textit{and} \quad r(b;p)=1-\varepsilon(p)-w(p)-\frac{c(b;p)}{\delta(2p-1)Z}$$

for all $p>\hat{p}$,

and

(12) $$r(a;p)=\varepsilon(p)+w(p)-\frac{c(a;p)}{\delta(2p-1)Z} \quad \textit{and} \quad r(b;p)=\varepsilon(p) \quad \textit{for all} \quad p<\hat{p}.$$

Proof: See Appendix 5.

From Theorem 2, the behavioral model $(\varepsilon(p),c(a,p),c(b;p))_{p\in(\underline{p},1)}$ is uniquely determined: for every $p>\hat{p}$

$\varepsilon(p)=1-r(a;p)$,

$c(a;p)=0$, and

$c(b;p)=\delta(2p-1)Z\{r(a;p)-r(b;p)-w(p)\}$,

and for every $p<\hat{p}$,

$\varepsilon(p)=r(b;p)$,

$c(a;p)=\delta(2p-1)Z\{w(p)-r(a;p)+r(b;p)\}$, and

$c(b;p)=0$.

Note that

(iii) the player is null-reciprocal at the critical level $\hat{p}$, that is, $c(a;\hat{p})=c(b;\hat{p})=0$,

(iv) $\varepsilon(p)$ is single-peaked with the peak at the critical level $\hat{p}$, that is, increasing in $p\in(\underline{p},\hat{p})$ and decreasing in $p\in(\hat{p},1)$,

(v) $c(a;p)$ is decreasing in $p\in(\underline{p},\hat{p})$,

and

(vi) $c(b;p)$ is increasing in $p\in(\hat{p},1)$.

9.3. Trade-Offs

The behavioral model $(\varepsilon(p),c(a,p),c(b;p))_{p\in(\underline{p},1)}$ is said to be *more kind* in p than in p' if either

it is more positively reciprocal in p than in p', that is, $c(a;p)>c(a;p')$, or

it is less negatively reciprocal in p than in p', that is, $c(b;p') > c(b;p)$.

From Theorem 2, the behavioral model has the following *trade-off between kindness and accuracy*:

(vii) the less kind a player is, the more accurate the monitoring technology is.

Given a sufficient level of monitoring accuracy, players tend to be more negatively reciprocal as monitoring is more accurate. This tendency makes the retaliation intensity more severe, and therefore, works against the better success in cooperation caused by the improvement of monitoring technology. Given an insufficient level of monitoring accuracy, players tend to be more positively reciprocal as monitoring is less accurate. This tendency makes the retaliation intensity milder, and thereby mitigates the worse success in cooperation caused by the deterioration of the monitoring technology.[29]

From Theorem 2, the behavioral model has the following *trade-off between naïveté and reciprocity*:

(viii) the more naively a player makes action choices, the less reciprocal he is.

When a player is negatively reciprocal, he tends to be more conscious, that is, less likely to select the defective action mistakenly, despite observing the good signal, as he is more negatively reciprocal. When a player is positively reciprocal, he tends to be more conscious, that is, less likely to mistakenly select the cooperative action despite observing the bad signal, as he is more positively reciprocal.

9.4. Uniqueness of G-TFT Equilibrium

We further show that the accuracy-contingent g-TFT strategy $(r(a;p), r(b;p))_{p\in(\underline{p},1)}$ is the only plausible accuracy-contingent g-TFT equilibrium in the derived behavioral

[29] The model here and that in Duffy and Muñoz-García (2012) share the common characteristic that social preferences play an important role in people undertaking collusion in repeated games. However, the implications of the social preferences are different. Duffy and Muñoz-García (2012) demonstrated that, motivated by fairness, social preference facilitates the achievement of collusion even if the time discount factor is insufficient. By contrast, our behavioral theory shows that the monitoring technology is a crucial determinant of whether social preference helps collusion or not due to the trade-off between kindness and accuracy. Motivated by reciprocity, social preference facilitates collusion even under poorer monitoring accuracies, in which a larger discount factor is required. However, social preference inhibits people from collude under better monitoring accuracies.

model $(\varepsilon(p), c(a,p), c(b;p))_{p\in(\underline{p},1)}$. In fact, whenever $\underline{p} < \hat{p}$, then there is no other accuracy-contingent g-TFT equilibrium strategy at all.

If $\underline{p} = \hat{p}$, there is another accuracy-contingent g-TFT strategy equilibrium, $(\tilde{r}(a;p), \tilde{r}(b;p))_{p\in(\underline{p},1)}$, which is specified as

$$\tilde{r}(a;p) = \varepsilon(p) + w(p) \text{ and } \tilde{r}(b;p) = \varepsilon(p) \text{ for all } p \in (\underline{p},1),$$

where $(\tilde{r}(a;p), \tilde{r}(b;p))_{p\in(\underline{p},1)}$ and $(r(a;p), r(b;p))_{p\in(\underline{p},1)}$ are the only equilibria in the derived behavioral model.

Note that $(\tilde{r}(a;p), \tilde{r}(b;p))_{p\in(\underline{p},1)}$ is quite implausible, because $\tilde{r}(a;p)$ and $\tilde{r}(b;p)$ are decreasing, and the frequency of cooperative action choices is, therefore, decreasing, at the level of monitoring accuracy. Note also that $(\tilde{r}(a;p), \tilde{r}(b;p))_{p\in(\underline{p},1)}$ is less efficient than $(r(a;p), r(b;p))_{p\in(\underline{p},1)}$, irrespective of $p \in (\underline{p},1)$.

The abovementioned uniqueness is in contrast to the standard theory, which has numerous g-TFT equilibria. The uniqueness of our theory substantially relies on the inclusion of psychological costs $c(a;p)$ and $c(b;p)$, which make the equilibrium constraints dependent on the signal occurrence in the previous round.

10. Conclusion

This study experimentally examines collusion in a repeated prisoner's dilemma game with random termination in which monitoring is imperfect and private. Each player obtains information about the opponent's action choice through a signal instead of a direct observation, and the signal the opponent observes is not observable to the player. We assume that the continuation probability is large enough to support collusion even with the poor monitoring technology. Our study is the first experimental attempt to investigate imperfect private monitoring.

Our experimental results indicate that a significant proportion of our subjects employed g-TFT strategies, which are straightforward stochastic extensions of the well-known TFT strategy. We significantly depart from the experimental literature by focusing on g-TFT strategies, which have attracted less attention in the empirical literature despite their theoretical importance. Our finding that a significant proportion of our subjects follow g-TFT strategies reveals its empirical importance.

Although many subjects follow g-TFT strategies, their retaliating policies systematically deviate from the predictions by the g-TFT equilibria. Our subjects retaliate more in the high accuracy treatment than in the low accuracy treatment, which is contrary to the theoretical implications. The subjects retaliate more in the high accuracy treatment, while they retaliate lesser in the low accuracy treatment. These experimental findings indicate that the subjects fail to improve their welfare by effectively utilizing the monitoring technology, as the standard theory predicts.

As feedback from these experimental findings to the theoretical development, we characterize a behavioral model of preferences that incorporates reciprocity and naïveté. Our behavioral theory describes signal-contingent behavior consistent with our experimental results as a unique, plausible g-TFT equilibrium. This feedback could help inform the development of a pervasive theory that has more relevance to real behavior and more predictive power.

References

Abreu, D. (1988): "On the Theory of Finitely Repeated Games with Discounting," *Econometrica* 56, 383–396.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041–1063.

Aoyagi, M., V. Bhaskar, and G. Fréchette (2015): "The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private," Discussion Paper No. 942, The Institute of Social and Economic Research, Osaka University.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior* 10, 122–142.

Breitmoser, Y. (2015): "Cooperation, but No Reciprocity: Individual Strategies in the Repeated Prisoner's Dilemma," *American Economic Review* 105(9), 2882–2910.

Dal Bó, P. (2005): "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review* 95, 1591–1604.

Dal Bó, P. and G. Fréchette (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review* 101(1), 411–429.

Dal Bó, P. and G. Fréchette (2014): "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," mimeo.

Duffy, J. and F. Muñoz-García (2012): "Patience or Fairness? Analyzing Social Preferences in Repeated Games," *Games* 3(1), 56–77.

Ely, J. and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory* 102, 84–105.

Fehr, E. and S. Gächter (2000): "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14, 159–181.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics* 10(2), 171–178.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997–1040.

Fudenberg, D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica* 54, 533–554.

Fudenberg, D., D.G. Rand, and A. Dreber (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review* 102(2), 720–749.

Green, E. and R. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica* 52, 87–100.

Goeree, J., C. Holt, and T. Palfrey (2008): "Quantal Response Equilibrium," *New Palgrave Dictionary of Economics*, Second Edition, Edited by S. Durlauf and L. Blume, Palgrave Macmillan.

Güth, W., R. Schmittberger, and B. Schwarze (1982): "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization* 3, 367–388.

Mailath, J. and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica* 72, 823–852.

Matsushima, H. (2013): "Interlinkage and Generous Tit-For-Tat Strategy," *Japanese Economic Review* 65, 116–121.

Matsushima, H., T. Tanaka, and T. Toyama (2013): "Behavioral Approach to Repeated Games with Private Monitoring," DP CIRJE-F-879, University of Tokyo.

Matsushima, H. and T. Toyama (2011): "Monitoring Accuracy and Retaliation in Infinitely Repeated Games with Imperfect Private Monitoring: Theory and Experiments", DP CIRJE-F-795, University of Tokyo.

Matsushima, H., T. Toyama, and N. Yagi (2007): "Repeated Games with Imperfect Monitoring: An Experimental Approach," University of Tokyo.

Nowak, M. and K. Sigmund (1992): "Tit-For-Tat in Heterogeneous Populations," *Nature* 355, 250–253.

Piccione, M. (2002): "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory* 102, 70–83.

Radner, R. (1986): "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies* 53, 43–57.

Sugaya, T. (2012): *The Folk Theorem in Repeated Games with Private Monitoring*, Ph. D. dissertation.

Takahashi, S. (2010): "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory* 145, 42–64.

Table 1:
Prisoner's Dilemma with Symmetry and Additive Separability
$(X, Y, Z) = (60, 10, 55)$

	A	B
A	60 60	5 70
B	70 5	15 15

Table 2:
Features of Experimental Design

	Number of subjects	Treatment order (sequence of game lengths)
October 5, 2006 (10:30–12:30)	28	0.6 (24, 40, 25), 0.9 (28, 33, 14)
October 5, 2006 (14:30–16:30)	24	0.6 (20, 23, 37), 0.9 (34, 34, 19)
October 6, 2006 (10:30–12:30)	28	0.9 (38, 21, 25), 0.6 (25, 28, 29)
October 6, 2006 (14:30–16:30)	28	0.9 (25, 35, 23), 0.6 (36, 30, 21)

Table 3:
Decisions and Signal

	$p = 0.9$			$p = 0.6$		
	N	Mean	St. Dev.	N	Mean	St. Dev.
Cooperative choice	8,864	0.672	0.469	9,144	0.355	0.479
Good signal	8,864	0.637	0.481	9,144	0.483	0.500

Table 4:
Means of Cooperative Action Choice

	$p = 0.9$	$p = 0.6$	p-values
$q(p)$ (round 1)	0.781 (0.042)	0.438 (0.035)	< 0.001
$r(a; p)$	0.852 (0.033)	0.437 (0.018)	<0.001
Individual-level means	0.799 (0.032)	0.423 (0.027)	< 0.001+
$r(b; p)$	0.344 (0.026)	0.272 (0.030)	< 0.001
Individual-level means	0.448 (0.026)	0.279 (0.032)	< 0.001+

Notes: The standard errors (shown in parentheses) are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values. The null hypothesis is that the values are identical across the two treatments.
+ The Wilcoxon matched-pair test within individual rejects the null hypothesis that the values are identical across the two treatments (p < 0.001 for each).

Table 5:
Retaliation Intensities

	Mean	*S.E.*	*p-value*
$r(a;0.9)-r(b;0.9)$	0.508	0.028	< 0.001+
Individual-level means	0.352	0.028	< 0.001+
$r(a;0.6)-r(b;0.6)$	0.165	0.025	< 0.001+
Individual-level means	0.144	0.023	< 0.001+
$(r(a;0.9)-r(b;0.9))-(r(a;0.6)-r(b;0.6))$	0.344	0.031	< 0.001
Individual-level means	0.208	0.029	< 0.001

Notes: The standard errors are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values.
+ The hypothesis tests for the comparison to the value implied by the standard theory (w(p)), which is 0.235 in the high accuracy treatment and 0.94 in the low accuracy treatment. The null hypothesis is that the mean is identical to the implied value.

Table 6:
Reduced-form Regression Results on Signal Contingency

	p = 0.9	**p = 0.6**
aA	0.887*** (0.019)	0.766*** (0.038)
aB	0.502*** (0.040)	0.255*** (0.035)
bA	0.506*** (0.038)	0.582*** (0.049)
bB	0.106*** (0.021)	0.110*** (0.042)
aAa	0.062*** (0.022)	0.037 (0.031)
aBa	0.017 (0.054)	-0.055 (0.034)
bAa	0.125*** (0.041)	-0.045 (0.046)
bBa	0.132*** (0.037)	0.022 (0.016)
Observations	8,216	8,496
R2	0.816	0.531
Adjusted R2	0.816	0.531

Notes: xYz in the regressors denote that the player takes action Y, observes signal x about the opponent's choice, and observed signal z in the previous round. Similarly, xY denotes that the player takes action Y and observes signal x about the opponent's choice. The standard errors (shown in parentheses) are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values. *p < 0.1, **p < 0.05, ***p < 0.01.

Table 7:
Strategy Set in our SFEM

Strategy	Description
All-C	Always cooperate
TFT	Tit-for-tat
g-TFT- $r(a)$ **-** $r(b)$	Generous tit-for-tat, cooperate if a good signal occurs with probability $r(a)$; forgive in the event of a bad signal and cooperate with probability $r(b)$
All-D	Always defect
TF2T	Tit-for-two-tat (retaliate if bad signals occur in all of the last two rounds)
g-TF2T- r	Generous tit-for-two-tat, playing cooperation stochastically with probability r even after observing two consecutive bad signals
TF3T	Tit-for-three-tat (retaliate if a bad signal occurs in all of the last three rounds)
2TFT	Two tit-for-tat (retaliate twice consecutively if a bad signal occurs)
g-2TFT- r	Generous two tit-for-tat, certainly retaliate if a bad signal occurs, but forgive and cooperate with probability r in the next round if a good signal occurs (second punishment)
2TF2T	Two tit-for-two-tat (retaliate twice consecutively if a bad signal occurs in all of the last two rounds)
Grim	Cooperate until the player chooses defection or observes a bad signal, and then play defection forever
Grim-2	Cooperate until the case of twice in a row occurs, in which the player chooses defection or observes a bad signal, and then play defection forever
Grim-3	Cooperate until the case of three times in a row occurs, in which the player chooses defection or observes a bad signal, and then play defection forever
Random- r	Cooperate with probability r irrespective of signals

Table 8:
Maximum Likelihood Estimates of Individual Strategies

	p = 0.9	S.E.	p = 0.6	S.E.
All-C (g-TFT-1-1)	0	(0.029)	0.037**	(0.019)
g-TFT-1-0.875	0	(0.009)	0	(0)
g-TFT-1-0.75	0	(0)	0	(0.002)
g-TFT-1-0.625	0	(0.040)	0	(0)
g-TFT-1-0.5	0.171***	(0.063)	0	(0)
g-TFT-1-0.375	0	(0)	0	(0)
g-TFT-1-0.25	0.041	(0.031)	0	(0)
g-TFT-1-0.125	0	(0.007)	0.010	(0.014)
TFT (g-TFT-1-0)	0.026	(0.016)	0.027	(0.022)
Ramdom-0.875 (g-TFT-0.875-0.875)	0.002	(0.014)	0	(0)
g-TFT-0.875-0.75	0.033	(0.024)	0.027	(0.017)
g-TFT-0.875-0.625	0.022	(0.033)	0	(0.015)
g-TFT-0.875-0.5	0.042	(0.041)	0.004	(0.015)
g-TFT-0.875-0.375	0.055	(0.048)	0	(0)
g-TFT-0.875-0.25	0.107***	(0.041)	0	(0)
g-TFT-0.875-0.125	0	(0.006)	0.010	(0.011)
g-TFT-0.875-0	0	(0)	0.010	(0.009)
Random-0.75 (g-TFT-0.75-0.75)	0	(0)	0	(0)
g-TFT-0.75-0.625	0	(0.001)	0.038	(0.024)
g-TFT-0.75-0.5	0	(0.017)	0.017	(0.022)
g-TFT-0.75-0.375	0.003	(0.023)	0.026	(0.027)
g-TFT-0.75-0.25	0	(0)	0	(0.004)
g-TFT-0.75-0.125	0.011	(0.015)	0	(0)
g-TFT-0.75-0	0.009	(0.008)	0	(0)
Random-0.625 (g-TFT-0.625-0.625)	0.011	(0.010)	0.014	(0.018)
g-TFT-0.625-0.5	0	(0.005)	0	(0.012)
g-TFT-0.625-0.375	0.023	(0.029)	0.089*	(0.046)
g-TFT-0.625-0.25	0	(0)	0	(0.015)
g-TFT-0.625-0.125	0.026	(0.021)	0	(0)
g-TFT-0.625-0	0	(0.010)	0	(0)
Random-0.5 (g-TFT-0.5-0.5)	0	(0.013)	0.023	(0.024)
g-TFT-0.5-0.375	0.053	(0.034)	0.050	(0.045)
g-TFT-0.5-0.25	0.003	(0.011)	0.038	(0.029)
g-TFT-0.5-0.125	0	(0)	0	(0.006)
g-TFT-0.5-0	0.019	(0.014)	0	(0)
Random-0.375 (g-TFT-0.375-0.375)	0.021	(0.019)	0.044	(0.045)
g-TFT-0.375-0.25	0	(0.006)	0.090	(0.060)
g-TFT-0.375-0.125	0	(0)	0	(0.012)
g-TFT-0.375-0	0	(0.003)	0.014	(0.013)
Random-0.25 (g-TFT-0.25-0.25)	0	(0)	0.063	(0.053)
g-TFT-0.25-0.125	0	(0)	0.069*	(0.041)
g-TFT-0.25-0	0.008	(0.014)	0.001	(0.011)
Random-0.125 (g-TFT-0.125-0.125)	0	(0.001)	0.033	(0.028)
g-TFT-0.125-0	0.029	(0.021)	0.037	(0.022)
All-D (g-TFT-0-0)	0.028	(0.017)	0.190***	(0.035)
g-TF2T-0.875	0.074*	(0.040)	0	(0)
g-TF2T-0.75	0	(0.021)	0	(0)
g-TF2T-0.625	0	(0.010)	0	(0.003)
g-TF2T-0.50	0.022	(0.028)	0	(0)
g-TF2T-0.375	0	(0.020)	0	(0)
g-TF2T-0.25	0	(0.024)	0.009	(0.009)
g-TF2T-0.125	0	(0.001)	0	(0)

TF2T (g-TF2T-0)	0	(0.004)	0	(0)
TF3T	0.070	(0.045)	0	(0.004)
g-2TFT-0.875	0	(0)	0	(0)
g-2TFT-0.75	0.009	(0.016)	0	(0)
g-2TFT-0.625	0	(0.024)	0	(0)
g-2TFT-0.50	0	(0.005)	0	(0)
g-2TFT-0.375	0	(0)	0	(0)
g-2TFT-0.25	0	(0)	0	(0)
g-2TFT-0.125	0	(0)	0	(0)
2TFT (g-2TFT-0)	0	(0)	0	(0)
2TF2T	0	(0.013)	0.009	(0.010)
Grim	0.009	(0.008)	0	(0)
Grim-2	0.013	(0.016)	0	(0.007)
Grim-3	0.035	(0.025)	0.020	(0.024)
Gamma	0.280***	(0.035)	0.247***	(0.059)

Notes: The standard errors (shown in parentheses) are cluster-bootstrapped (individual-level) with 100 repetitions, which is used to calculate p-values. *p < 0.1, **p < 0.05, ***p < 0.01.

Table 9:
Frequency of Cooperative Strategies, g-TFT, Retaliation Intensities (RI)

		p = 0.9		p = 0.6
Cooperative strategies		83.9%		34.7%
Non-cooperative strategies		16.1%		65.3%
Family of g-TFT		70.6%		55.6%
Including signal non-contingent, zero RI strategies (All-C, All-D, Random-r)		76.8%		96.2%
Including multi-round punishing strategies (g-2TFT)		77.7%		96.2%
Equilibrium RI	(RI = 0.25)	5.6%	(RI = 1)	2.7%
Stronger RI	(RI > 0.25)	54.4%	(RI > 1)	0.0%
Weaker RI (positive RI only)	(0 < RI < 0.25)	11.4%	(0 < RI < 1)	53.0%
Including zero RI strategies	(RI < 0.25)	17.7%	(RI < 1)	93.5%
Mean RI conditional on g-TFT		0.426		0.148
Median RI conditional on g-TFT		0.5		0.125
Lenient strategies		21.4%		3.7%

For Online Publication

Appendix 1: Screens in Constant Random Termination

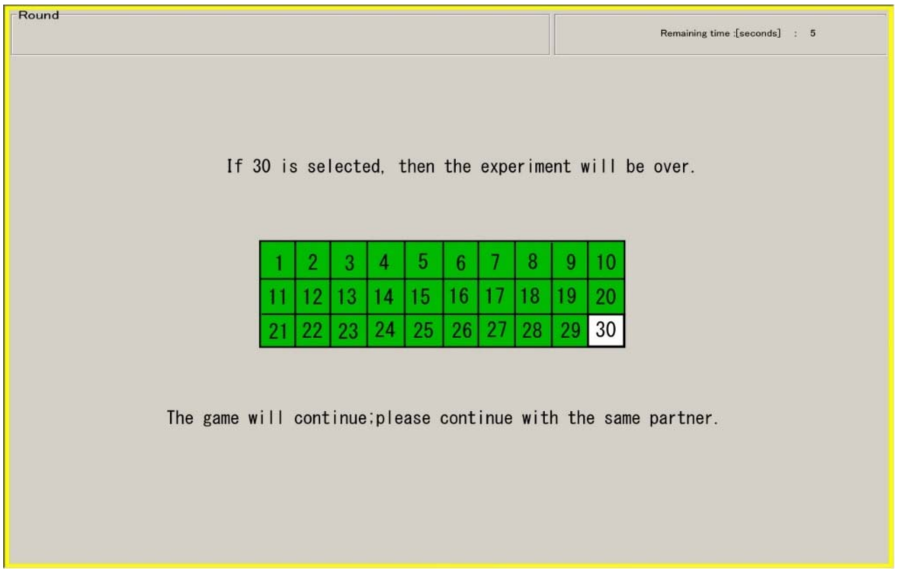


Figure A.1: Screen when the repeated game continues

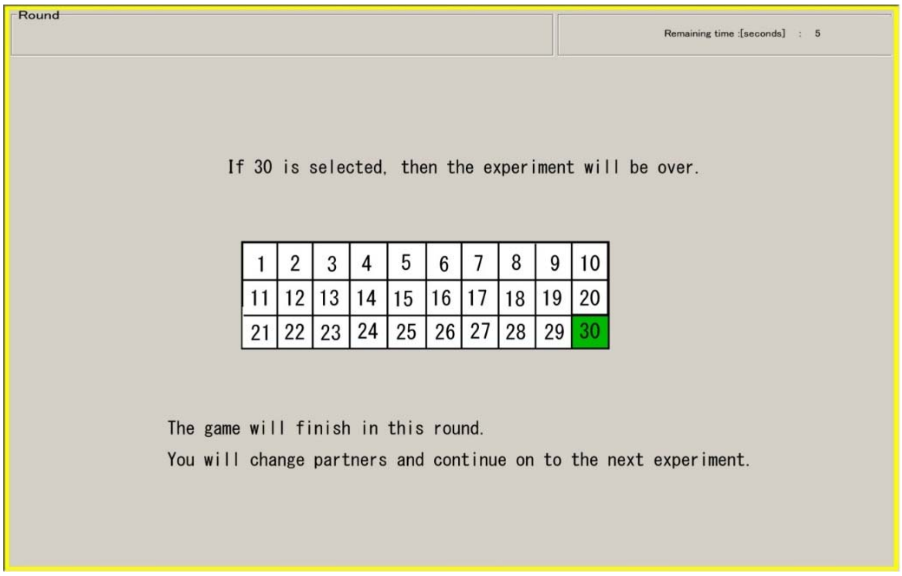


Figure A.2: Screen when the repeated game is terminated

Appendix 2: Impact of Experience

Several existing studies have reported that the frequency of cooperation changes as people experience playing repeated games (see discussion in Dal Bó and Fréchette, 2014). The overall cooperation rates might rise with experience of repeated games when the experimental parameters are conductive for cooperation. Since the welfare is uniquely determined by the overall cooperation rates in our payoff setting (i.e., the additive separability payoff structure), a shift of the overall cooperation rates implies a shift of welfare of the two players.

To examine the impact of the experience of repeated games on overall cooperation rates, we perform the following a reduced-form, linear regression analysis; we regress the action choices on the two explanatory variables, *RG2* and *RG3*. The dummy variable *RG2* (*RG3*) takes 1 if the choice is made in the second (third) repeated games in each treatment. Moreover, to further examine the change of the cooperation rate within a repeated game, we include the dummy variable *First 14 Rd*, which takes 1 if the choice is made in the first 14 rounds of each repeated game. The regression model is a fixed-effect model in which the individual heterogeneity in the tendency to adopt cooperative choices is controlled by individual fixed effects.

[TABLE A. 1 HERE]

Table A.1 displays the regression results. In the high accuracy treatment, the coefficient on the second repeated game is 0.029 (insignificant), and that on the third repeated game is 0.105 (significant, $p < 0.001$). These positive values in the regression coefficients indicate that our subjects tend to cooperate more as they gain experience, indicating that the welfare of the two players is improved by experience. However, the size is at most 11%, which is only a marginal effect.

In the low accuracy treatment, the coefficient on the second repeated game is −0.071 (significant, $p < 0.05$), and that on the third repeated game is −0.118 (significant, $p < 0.001$). Contrary to the case in the high accuracy treatment, our subjects tend to become less cooperative as they gain experience, even in the situation in which the experimental parameters are conductive for cooperation, rather indicating that welfare worsens with experience. However, the sizes of the effect of experience are at most 12%, which is again only a marginal effect.

As for the effect within a repeated game, the coefficients on the first 14 rounds are statistically significantly and larger than 0 in both treatments ($p < 0.001$ for both treatments), although the sizes are at most 8%. Our subjects tend to become less cooperative as the rounds proceed in each repeated game, but the effect is small.

These results indicate that, although there are some experience effects on action choices, the sizes of the effects are not remarkably large in our data. In addition, we perform an identical analysis to the signal contingent cooperation rates, and find qualitatively similar results (Tables A.2 and A.3).

[TABLE A.2 HERE]

[TABLE A.3 HERE]

In addition to the cooperation rate, we investigate the effect of experience on retaliation intensities. Here, we perform a similar, reduced-form regression analysis, regressing the action choices on the dummy variable *Signal*, which takes 1 if the signal is good. The coefficient on the dummy variable captures the contrast between the cooperation rate contingent on the good signal and that on bad signal, which is the retaliation intensity. To examine the experience effects on retaliation intensities across repeated games and within a repeated game, we add the cross-product terms with *RG2*, *RG3*, and *First 14 Rd* in the set of the explanatory variables. Again, the regression model is a fixed-effect model in which the individual heterogeneity in the tendencies to adopt cooperative choices is controlled by individual fixed effects.

[TABLE A.4 HERE]

Table A.4 displays the regression results. None of the coefficients on the joint effect of *Signal* and repeated games (*RG2* and *RG3*) is significantly different from 0 in both treatments, implying that the retaliation intensities do not differ either in the second repeated games or in the third repeated games from that in the first repeated games.

Furthermore, the coefficient on the cross-product of *Signal* and *RG2* does not differ from that on the cross-product of *Signal* and *RG3* in both treatments (p = 0.702 for the high accuracy treatment, and p = 0.672 for the low accuracy treatment). These results jointly indicate that the retaliation intensities are stable across repeated games.

As for the within repeated game difference, only the coefficient on the joint effect of Signal and First 14 Rd in the high accuracy treatment significantly differs from 0, although the size of the effect is approximately 5%. Our subjects tend to rely on stronger retaliation intensities as the rounds proceed in a repeated game, but the difference is not remarkably large.

Overall, the results here indicate that the retaliation intensities do not change remarkably as our subjects gain experience.

Table A.1:
Fixed-effect Model Regression Results on the Experience Effect on Overall Cooperation Rates

	p = 0.9	**p = 0.6**
RG2	0.029 (0.031)	-0.071** (0.027)
RG3	0.105*** (0.031)	-0.118*** (0.029)
First 14 Rd	0.062*** (0.014)	0.080*** (0.015)
Observations	8,864	9,144
R2	0.021	0.025

Notes: Cluster-robust (individual-level) standard errors in parenthesis. *p < 0.1, **p < 0.05, ***p < 0.01. The coefficient on *RG3* is significantly larger than that on *RG2* in the high accuracy treatment (F-test, p = 0.016). The coefficient on *RG3* is significantly smaller than that on *RG2* in the low accuracy treatment (F-test, p = 0.027).

Table A.2:
Fixed-effect Model Regression Results on the Experience Effect on Cooperation Rates Contingent on Good Signals

	p = 0.9	**p = 0.6**
RG2	0.014 (0.016)	-0.055* (0.029)
RG3	0.076*** (0.021)	-0.104*** (0.034)
First 14 Rd	0.018 (0.011)	0.073*** (0.018)
Observations	5,453	4,265
R2	0.005	0.021

Notes: Cluster-robust (individual-level) standard errors in parenthesis. *p < 0.1, **p < 0.05, ***p < 0.01.

Table A.3:
Fixed-Effect Model Regression Results on the Experience Effect on Cooperation Rates Contingent on Bad Signals

	p = 0.9	p = 0.6
RG2	0.025 (0.030)	-0.070** (0.030)
RG3	0.052 (0.036)	-0.123*** (0.029)
First 14 Rd	0.061*** (0.018)	0.070*** (0.016)
Observations	3,087	4,555
R2	0.009	0.027

Notes: Cluster-robust (individual-level) standard errors in parenthesis. *p < 0.1, **p < 0.05, ***p < 0.01.

Table A.4:
Fixed-effect Model Regression Results on the Experience Effect on Retaliation Intensities

	p = 0.9	p = 0.6
Signal	0.392*** (0.030)	0.118*** (0.025)
Signal: RG2	-0.015 (0.027)	0.020 (0.022)
Signal: RG3	0.007 (0.038)	0.034 (0.027)
Signal: First 14 Rd	-0.048** (0.019)	0.005 (0.018)
RG2	0.029 (0.034)	-0.077** (0.030)
RG3	0.070* (0.039)	-0.130*** (0.029)
First 14 Rd	0.068*** (0.018)	0.072*** (0.016)
Observations	8,540	8,820
R2	0.021	0.025

Notes: Cluster-robust (individual-level) standard errors in parenthesis. *p < 0.1, **p < 0.05, ***p < 0.01. The coefficient on the cross-product term, *Signal*: *RG2*, is not significantly different from that on the cross-product term, *Signal*: *RG3*, in the high accuracy treatment (F-test, p = 0.702). The coefficient on the cross-product term, *Signal*: *RG2*, is not significantly different from that on the cross-product term, *Signal*: *RG3*, in the low accuracy treatment (F-test, p = 0.672).

Appendix 3: Derivation of Likelihood

The likelihood function in SFEM frameworks is derived as follows. The choice probability of subject i employing strategy s in round r of repeated game k, given the history of her past choices and signals obtained from the opponent up to the round, is defined as

(A. 1) $$P_{ikr}(s) = \frac{1}{1+\exp(-1/\gamma)},$$

if the observed choice is matched with the predicted choice by strategy s given the history up to the round. Otherwise, the choice is classified as implementation error and the choice probability is

(A. 2)
$$P_{ikr}(s) = \frac{1}{1+\exp(1/\gamma)},$$

where γ captures the probability of the implementation error.

The likelihood of subject i employing strategy s is

$$P_i(s) = \prod_k \prod_r P_{ikr}(s).$$

In the SFEM framework, the likelihood of subject i over all strategies is a finite mixture of $P_i(s)$ over the entire strategy set. We denote the frequency of occurrence of strategy s by $P(s)$. Then, the log likelihood of the MLE is

$$LH = \sum_i \ln\left(\sum_s P(s)P_i(s)\right).$$

The choice probabilities in (A. 1) and (A. 2) are defined over deterministic strategies. Since our list of strategies considered in our SEFM (Table 7) includes stochastic strategies, the choice probabilities should be extended to cover stochastic cases. Following Fudenberg, Rand, and Dreber (2012), the choice probabilities (A. 1) and (A. 2) are extended to cover stochastic strategies, as follows,

(A. 3)
$$P_{ikr}(s) = s_{ikr}\left(\frac{1}{1+\exp(-1/\gamma)}\right) + (1-s_{ikr})\left(\frac{1}{1+\exp(1/\gamma)}\right)$$

if the observed choice is A,

(A. 4)
$$P_{ikr}(s) = (1-s_{ikr})\left(\frac{1}{1+\exp(-1/\gamma)}\right) + s_{ikr}\left(\frac{1}{1+\exp(1/\gamma)}\right)$$

if the observed choice is B,

where s_{ikr} is the probability of playing A in stochastic strategy s given the history up to the round. Observe that the new formulations of the choice probabilities (A. 3) and (A. 4) are reduced to the previous definition (A. 1) and (A. 2) when s_{ikr} takes either 1 or 0 as deterministic choices.

The standard error of the MLE is computed through a cluster-bootstrap (subject-level) with 100 resamples, which is also used to perform the hypothesis tests presented in Section 8.

Appendix 4: Robustness Checks of Our Strategy Estimation

In this part of the appendix, we discuss the robustness of the SFEM estimates. In the main text, we use all three repeated games of each treatment in our estimation. Here, we demonstrate that the estimation results show almost no changes even using only the final two repeated games in each treatment (Tables A.5 and A.6).

The mean retaliation intensity in the high accuracy treatment decreases slightly in the final two repeated games, which is closer to the value implied by the g-TFT equilibria (0.235), which might suggest that our subjects learn optimal retaliation intensities by experience in the high accuracy treatment.

Table A.5:
Aggregated Estimates for High Accuracy Case (p = 0.9)

	All	Final 2
Cooperative strategies	83.9%	88.3%
Family of g-TFT (including zero RI strategies)	76.8%	75.2%
Equilibrium RI	5.6%	7.6%
Stronger RI	54.4%	49.0%
Weaker RI (including zero RI strategies)	17.7%	20.8%
Mean RI conditional on g-TFT (including zero RI strategies)	0.426	0.360
Lenient strategies	21.4%	21.6%

Table A.6:
Aggregated Estimates for Low Accuracy Case (p = 0.6)

	All	Final 2
Cooperative strategies	34.7%	37.6%
Family of g-TFT (including zero RI strategies)	96.2%	96.5%
Equilibrium RI	2.7%	2.4%
Stronger RI	0.0%	0.0%
Weaker RI (including zero RI strategies)	93.5%	94.0%
Mean RI conditional on g-TFT (including zero RI strategies)	0.148	0.135
Lenient strategies	3.7%	2.5%

Appendix 5: Proof of Theorem 2

The proof of the "if" part is straightforward from (7), (8), (9), (10), (11), and (12).

Fix an arbitrary $p \in (\underline{p}, 1)$. From continuity, we can assume without loss of generality that

$$1 - r(a; p) \neq r(b; p).$$

Suppose $p > \hat{p}$, that is,

$$r(a; p) - r(b; p) > w(p) > 0.$$

This, along with $w(p) \equiv \dfrac{Y}{\delta(2p-1)Z}$ and $c(a; p) \geq 0$, implies the inequality of (8), and therefore,

$$r(a; p) = 1 - \varepsilon(p) \text{ and } r(b; p) < 1 - \varepsilon(p).$$

Hence, from (9) and (10), either $r(b; p) = \varepsilon(p)$ or

$$Y + c(b; p) = \delta Z(2p - 1)\{r(a; p) - r(b; p)\},$$

which implies (11), where $r(b; p) \neq \varepsilon(p)$ holds because $r(a; p) = 1 - \varepsilon(p)$ and $1 - r(a; p) \neq r(b; p)$.

Next, suppose $p < \hat{p}$, that is,

$$r(a; p) - r(b; p) < w(p).$$

This, along with $w(p) \equiv \dfrac{Y}{\delta(2p-1)Z}$ and $c(b; p) \geq 0$, implies the inequality of (9), and therefore,

$$r(b; p) = \varepsilon(p) \text{ and } r(a; p) > \varepsilon(p).$$

Hence, from (7) and (8), either $r(a; p) = 1 - \varepsilon(p)$ or

$$Y - c(a; p) = \delta Z(2p - 1)\{r(a; p) - r(b; p)\},$$

which implies (12), where $r(a; p) \neq 1 - \varepsilon(p)$ holds because of $r(b; p) = \varepsilon(p)$ and $1 - r(a; p) \neq r(b; p)$.

Q.E.D.

Appendix 6: Experimental Instruction
(October 6, 2006, Translation from Japanese into English)

Please make sure all the contents are in your envelope. The envelope should have the following items.

1. Pen – 1
2. Instruction – 1 copy
3. Printed computer screen images – 1 copy
4. Bank transfer form – 1 sheet
5. Scratch paper – 1 sheet

If you have any missing item, please raise your hand quietly. We will collect the items at the end of all the experiments, except for the scratch paper, which you can keep.

Please look at the instructions (this material). You will be asked to make decisions at a computer terminal. You will earn "points" according to your performance in the experiments. The points will be converted into monetary rewards at the exchange rate of 0.6 yen per point, which will be paid in addition to the participation fee (1,500 yen). The total amount of money you will receive from the experiments is

the number of points earned × 0.6 yen + participation fee of 1,500 yen.

Any communication with other participants (i.e., conversation or exchange of signals) is not allowed during the experiments; otherwise, you may be asked to leave the experiments. Furthermore, you are not allowed to leave in the middle of the experiments unless an experimenter allows or asks you to do so. Please turn off your cell phones during the experiments.

Outline of Experiments

We will conduct six experiments across two sessions. Each session includes three experiments with one practice experiment preceding them. The six experiments are independent of each other; the records of one experiment are not transferred to the other experiments. The experiments are conducted via a computer network. You are asked to make decisions at a computer terminal and interact with other participants through the computer network.

All the participants will be divided into pairs in each experiment. The pairs are selected randomly by the computer.

Each experiment consists of several rounds (i.e., Rounds 1, 2, 3, etc.). Later, we will explain the rule that decides the number of rounds conducted in each experiment. In each round, you are asked to choose one among two alternatives, which will also be explained below.

Please raise your hand quietly if you have any questions.

Decision Making

You will be asked to choose either A or B in each round. Your partner will also be asked to choose either A or B. Please look at the table.

You \ Your partner	A		B	
A	60	60	5	70
B	70	5	15	15

The table summarizes the points you and your partner earn according to the combination of the choices made by the two players. The characters in the left column marked in red indicate your choice, which is either A or B. The characters in the top row marked in light blue indicate the choice of your partner, which is also either A or B. In each cell, the numbers in red on the left side indicate the points you earn, and the numbers in light blue on the right side indicate the points your partner earns.

If both you and your partner select A,

both you and your partner earn 60 points.

If you select A and your partner selects B,

you earn 5 points, and your partner earns 70 points.

If you select B and your partner selects A,

you earn 70 points, and your partner earns 5 points.

If both you and your partner select B,

both you and your partner earn 15 points.

Please look at the table carefully, and ensure that you understand how the points will be awarded to you and your partner according to the choices made by the two players. Your earnings depend not only on your choice, but also on the choice of your partner. Similarly, your partner's earnings depend on your choice as well as her own.

Please raise your hand quietly if you have any questions.

Session 1

Session 1 consists of three experiments, 1, 2, and 3. The three experiments follow identical rules and will be conducted consecutively.

Observable Information

You are not allowed to observe whether your partner selected A or B directly. However, you will receive signal a or signal b, which has information about your partner's choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to you.

If your partner selects A, you will receive

signal a with 90% chance and signal b with 10% chance.

If your partner selects B, you will receive

signal b with 90% chance and signal a with 10% chance.

In the same way, your partner will not know whether you have selected A or B. However, your partner will receive signal a or signal b, which has information about your choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to your partner.

If you select A, your partner will receive

signal a with 90% chance and signal b with 10% chance.

If you select B, your partner will receive

signal b with 90% chance and signal a with 10% chance.

The signal you receive and the signal your partner receives are decided independently and have no correlation. Furthermore, the computer determines the signals independently in each round.

We refer to the stochastic rules for this signal-generating process as **signaling with 90% accuracy.**

Please raise your hand quietly if you have any questions.

Number of Rounds

The number of rounds in each experiment will be determined randomly. At the end of each round, the computer will randomly select a number from 1 to 30 without replacement, so there is a 1/30 chance for any number being selected by the computer. The number selected by the computer is applied uniformly to all participants.

The experiment will be terminated when the number 30 is selected by chance.

The experiment will continue if any number other than 30 is selected. However, you will notice only that a number other than 30 is selected, instead of the specific number selected by the computer. Then, you will move on to the next round, and will be asked to make a decision faced with the same partner.

The probability that the experiment is terminated in each round remains the same, which is 1/30, regardless of the number of rounds (1, 2, 3, etc.). However, the maximum possible number of rounds in an experiment is experimentally controlled, which is 98.

When Experiment 1 is terminated, you proceed to Experiment 2 and you will be randomly paired with a new partner. When Experiment 2 is terminated, you proceed to Experiment 3 and again, you will be randomly paired with a new partner. Session 1 will be over when Experiment 3 is terminated.

Please raise your hand quietly if you have any questions.

Description of Screens and Operations for Computers

Please look at the booklet with printed computer screen images.

Please look at Screen 1 and Screen 2. Screen 1 displays the screen that will be presented to you during the decision phases. Screen 2 is the screen that will be presented to your partner during the decision phases. Please look at the top left portion of each screen, which indicates that the current round is Round 4. The left portion of Screen 1 displays the information available to you up to the round. The left portion of Screen 2 presented to your partner displays the information available to her up to the round.

You are asked to click with the mouse to select either “A” or “B” in the bottom right portion of the screen. Then, the selection will be confirmed by clicking the “OK” button right below the alternatives.

Next, please look at Screen 3 and Screen 4. Screen 3 presents the results to you. Screen 4 presents the results to your partner. The screens display the situation in which, in Round 4, both you and your partner chose A. Screen 3 shows you that, in Round 4, “your partner’s signal (accuracy: 90%) is b,” indicating to you that the signal you observe about the partner’s choice is “b”. On the other hand, Screen 4 shows your partner that, in Round 4, “your partner’s signal (accuracy: 90%) is a”, indicating to your partner that the signal your partner observes about your choice is “a.” Recall that your partner will observe signal a with a probability of 90% and will observe signal b with a probability of 10% when you choose “A.”

Then, we move on to the lottery screens. Please turn the page and look at Screen 5 and Screen 6, which display the lottery. Any number from 1 to 30 will be randomly selected with an identical probability of occurrence, which is 1/30. Then, a part of the cells turns green according to the number selected. If the number 30 is selected, the cell numbered 30 turns green and the message below explains that the current experiment is terminated.

Otherwise, Screen 5 is shown in which all the cells numbered from 1 to 29 turn green at once (you do not know which number is selected specifically), and the message below explains that the experiment continues with the same partner. Screen 6 is presented when the number 30 is selected and the cell numbered 30 turns green, indicating that the current experiment is terminated in that round. Again, please make sure that the experiment is terminated when the cell numbered 30 turns green.

Finally, please look at Screen 7. This screen is presented at the end of each experiment. The screen displays the total number of points you earned in the experiment, the average number of points per round, the total number of points your partner earned, and the average number of points per round of your partner. Then, you will be rematched with a new partner and move on to the next experiment.

Please raise your hand quietly if you have any questions.

Session 2

Please look at page 6 of your instruction. In Session 2, you will participate in three experiments, namely, Experiments 4, 5, and 6, with a practice experiment preceding to them.

The three experiments follow identical rules and will be conducted consecutively. Session 2 proceeds similarly to Session 1, but the signal accuracy of Session 2 is different from that of Session 1. Except for the signal accuracy, the two sessions are identical.

Observable Information

You are not allowed to observe directly whether your partner selected A or B. However, you will receive signal a or signal b, which has information about your partner's choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to you.

If your partner selects A, you will receive

signal a with 60% chance and signal b with 40% chance.

If your partner selects B, you will receive

signal b with 60% chance and signal a with 40% chance.

In the same way, your partner will not know whether you selected A or B. However, your partner will also receive signal a or signal b, which has information about your choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to your partner.

If you select A, your partner will receive

signal a with 60% chance and signal b with 40% chance.

If you select B, your partner will receive

signal b with 60% chance and signal a with 40% chance.

The signals you and your partner receive are decided independently and have no correlation. Furthermore, the computer determines the signals independently in each round.

We refer to the stochastic rules for the signal-generating process as **signaling with 60% accuracy.**

Please raise your hand quietly if you have any questions.

Description of Screens and Operations for Computers

Please look at the booklet with printed computer screen images.

Please look at Screen 8 and Screen 9. Screen 8 displays the screen that will be presented to you during the decision phases. The left portion of Screen 8 displays the information available to you up to the round. Please check the current signal accuracy with the message "Signal accuracy for your partner's selection: 60%." Screen 9 is the screen that will be presented to your partner during the decision phases. The left portion of Screen 9 presented to your partner displays the information available to her up to the round.

Please look at Screen 10 and Screen 11 on page 6. Screen 10 presents the results to you. Screen 11 presents the results to your partner. The screens capture the situation in which, in Round 4, both you and your partner chose A, but you observed signal b and your partner observed signal a. Please make sure that the bottom right portions of Screen 10 or Screen 11 display the signals you and your partner observed, respectively.

The choices you made and the signals you observe about your partner's choices are only available to you, and are not available to your partner. Please make sure of this point in Screen 10 and Screen 11.

The screens for the lottery on the continuation of the experiment are identical to the case of Session 1, which shows numbers from 1 to 30. Please refer back to Screen 5 and Screen 6. The results screen at the end of the experiment is also identical to that in Session 1 (Screen 7).

Please raise your hand quietly if you have any questions.

Now, all the processes of the experiments have been completed, and all the points awarded to everyone recorded on the computer.

Please answer the questionnaire that will be distributed now.

Take the bank transfer form out of the envelope and fill it out accurately; otherwise, we will not be able to process the payment correctly for you.
Please raise your hand quietly if you have any questions.

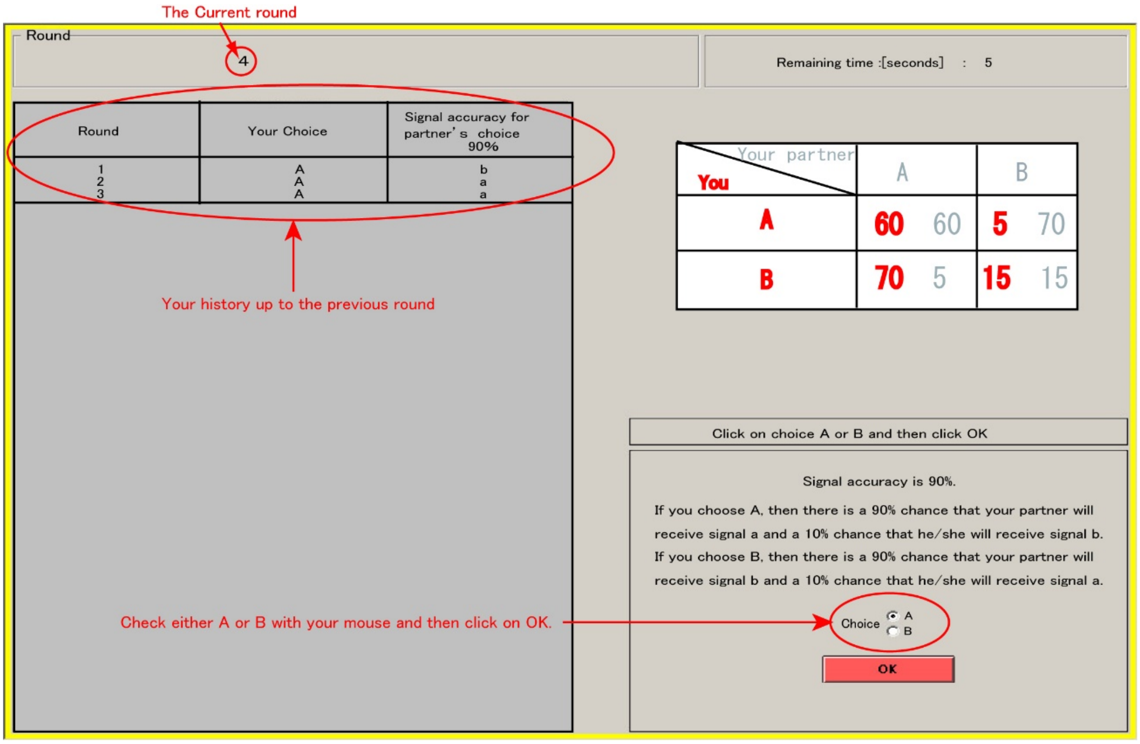
Please make sure that you fill out the questionnaire and the bank transfer form correctly.
Please raise your hand quietly if you have any questions.

Please put all the documents in the envelope. Please leave the pen and ink pad on the desk. Make sure you take all your belongings with you when you leave.

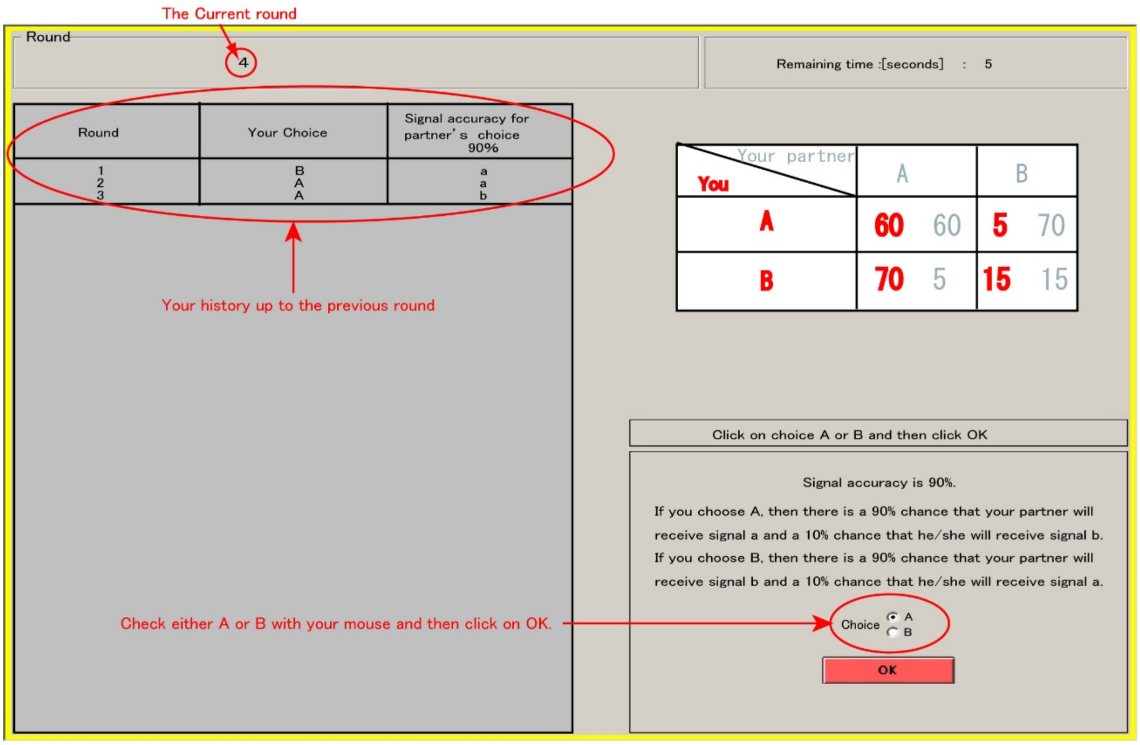
Please do not disclose any details regarding the experiments to anyone until Saturday. Thank you very much for your participation. Please follow the instructions of the experimenters to leave the room.

2. Computer Screen Images
(October 6, 2006, Translation from Japanese into English)

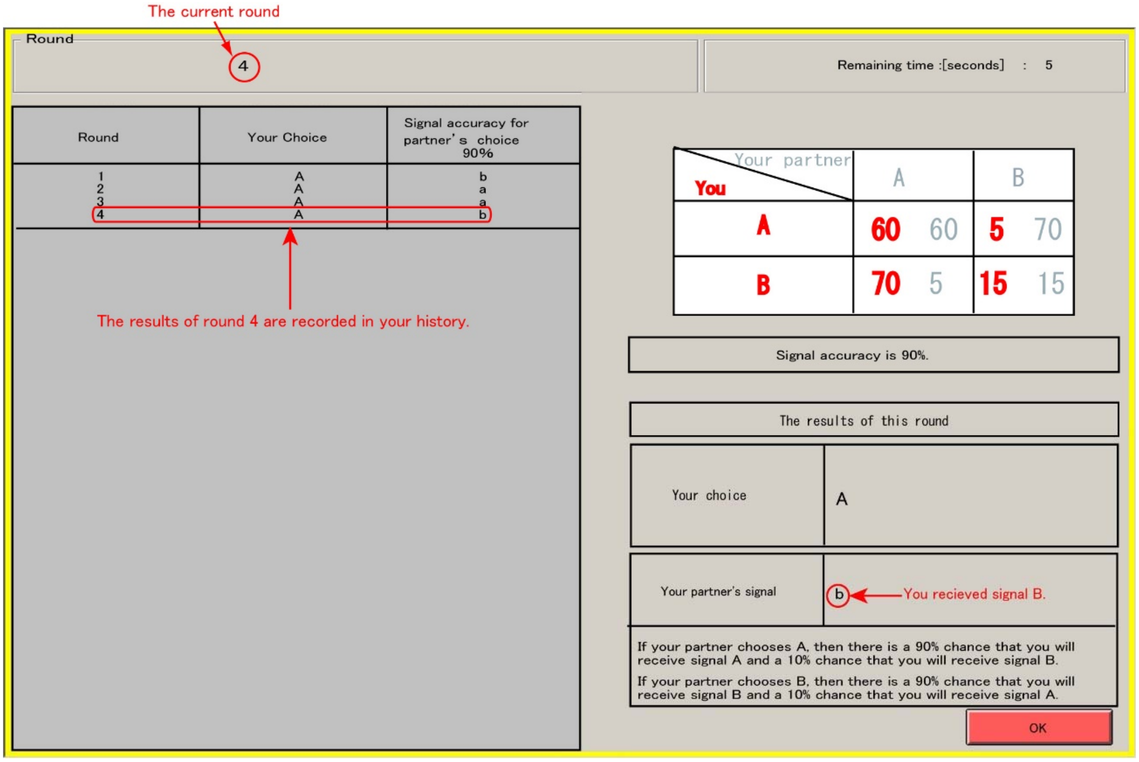
Screen 1: Your Selection Screen



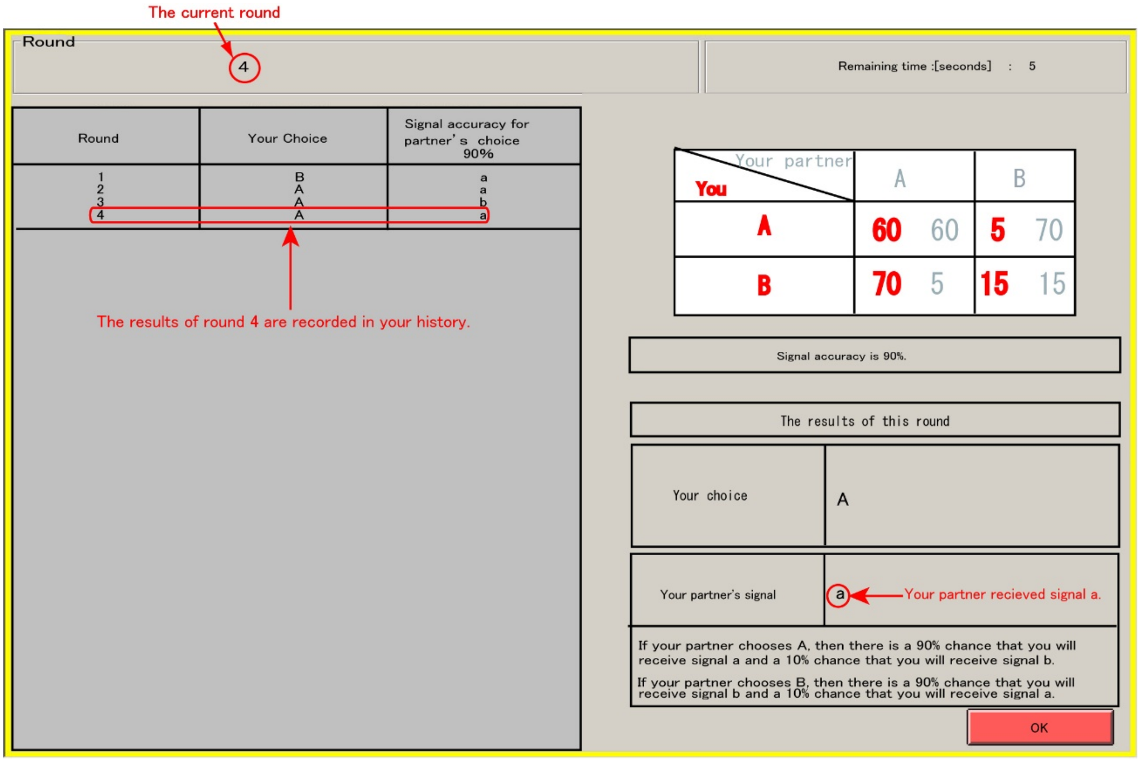
Screen 2: The Selection Screen of Your Partner



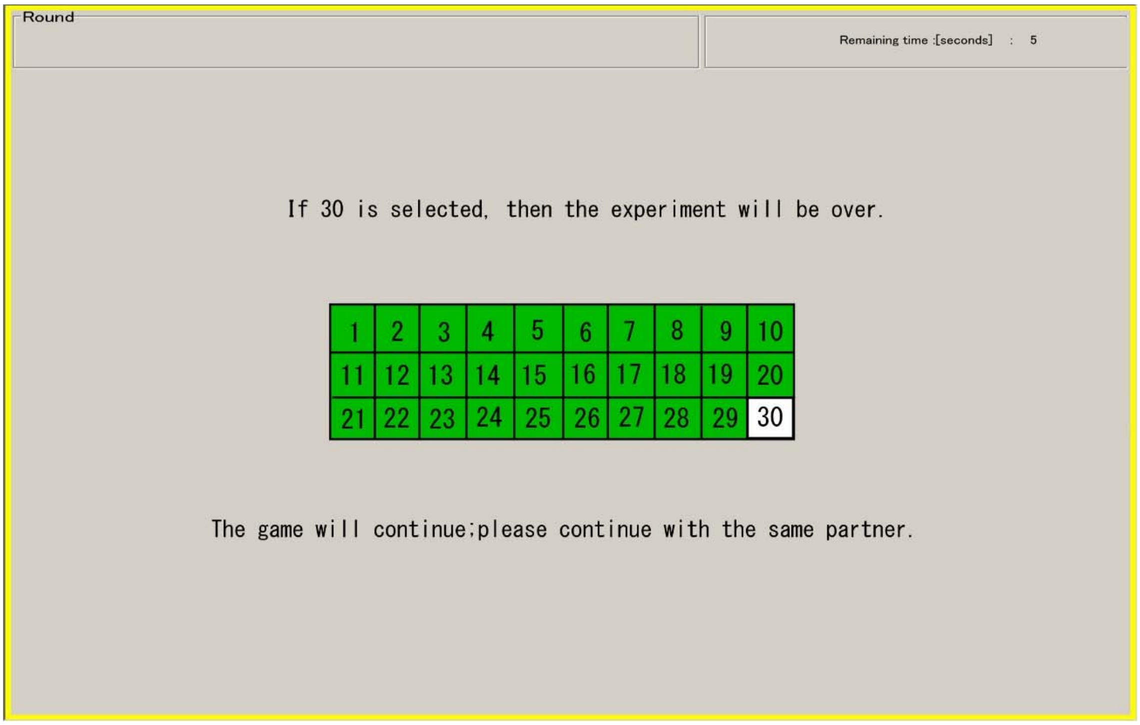
Screen 3: Your Results Screen



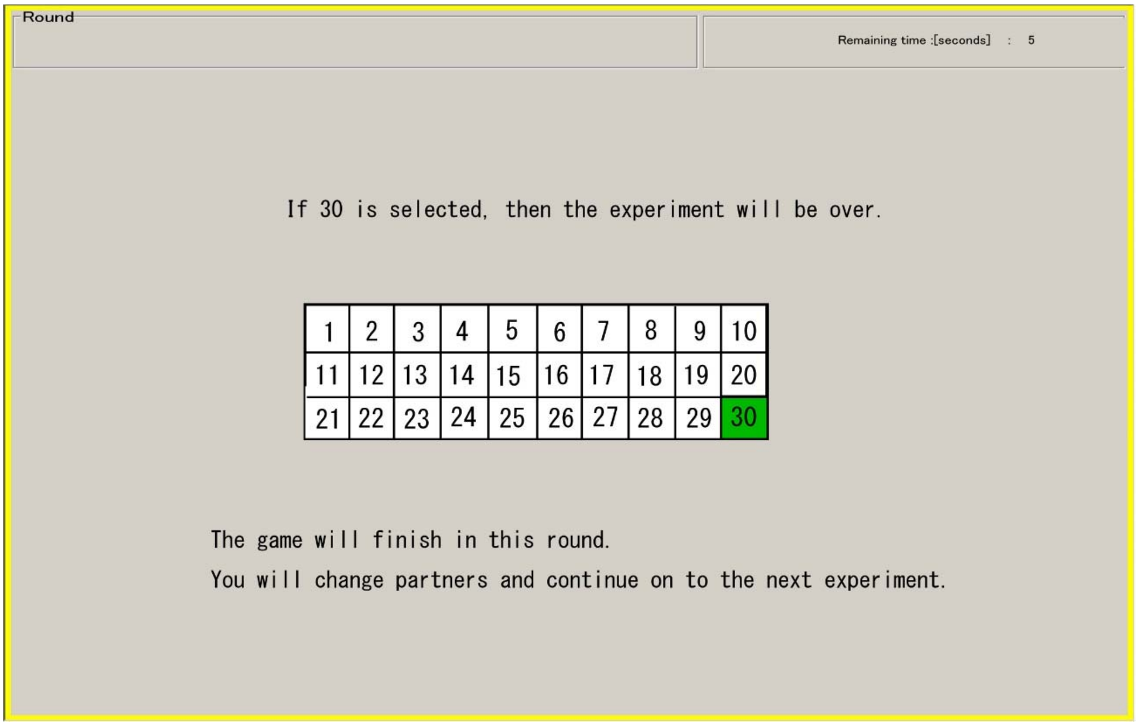
Screen 4: The Results Screen of Your Partner



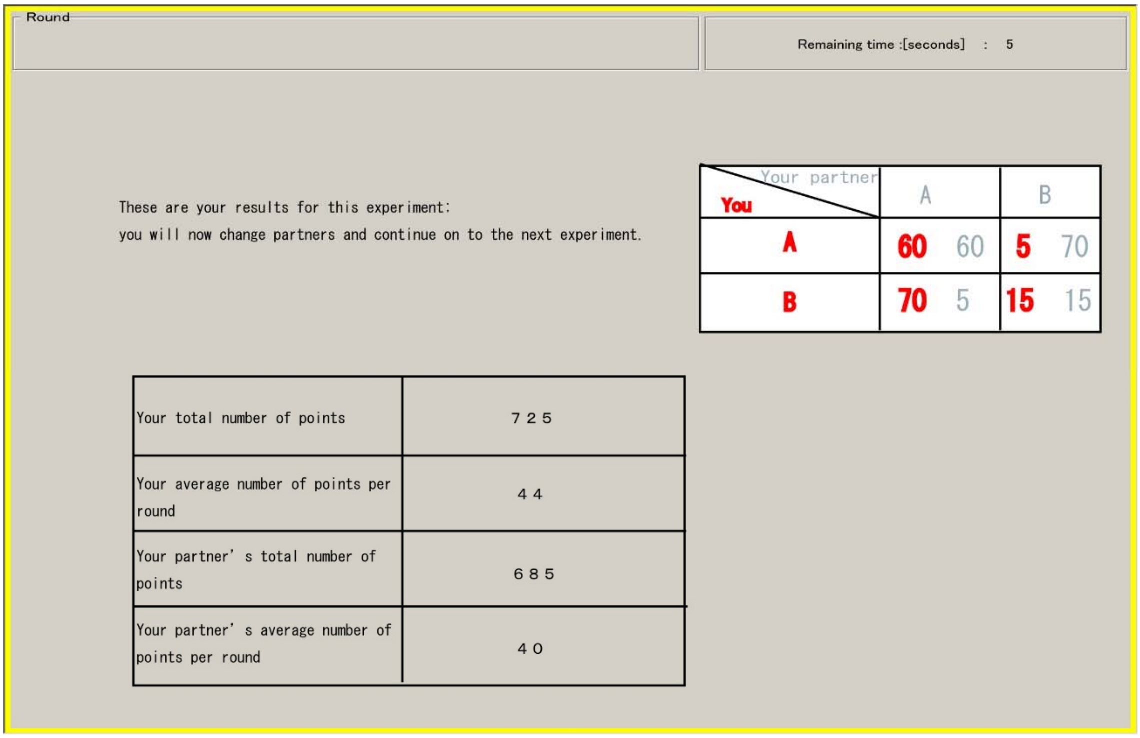
Screen 5: Lottery (experiment continues)



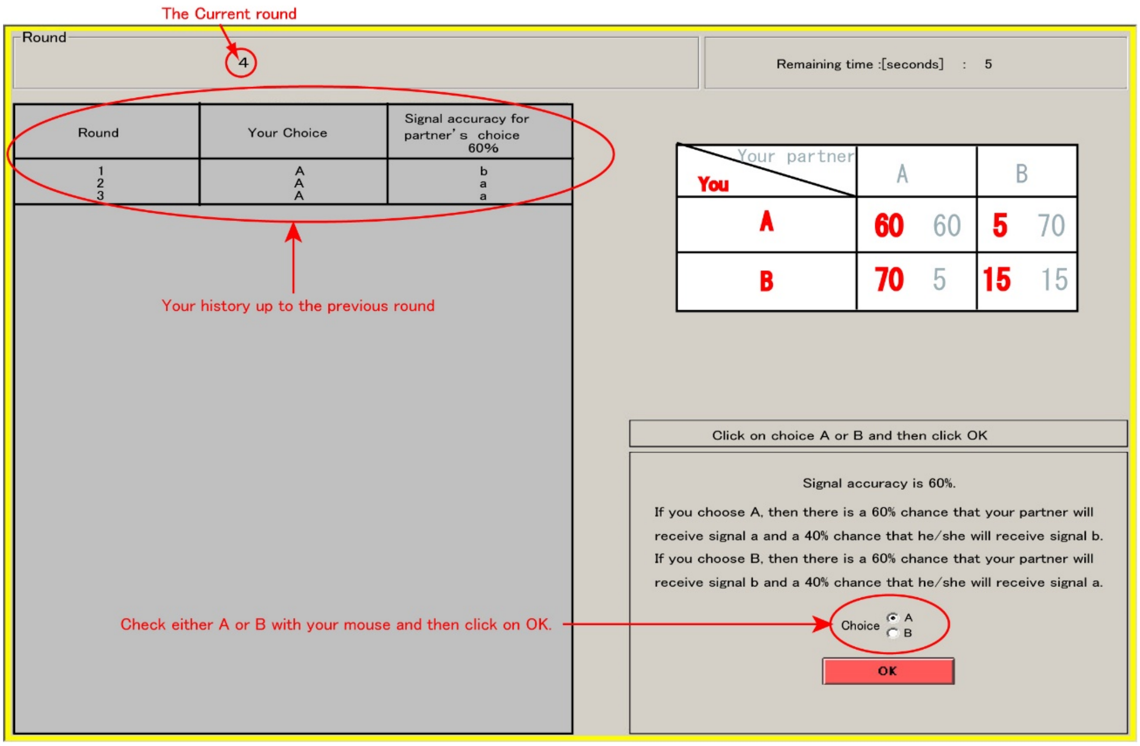
Screen 6: Lottery (experiment is over)



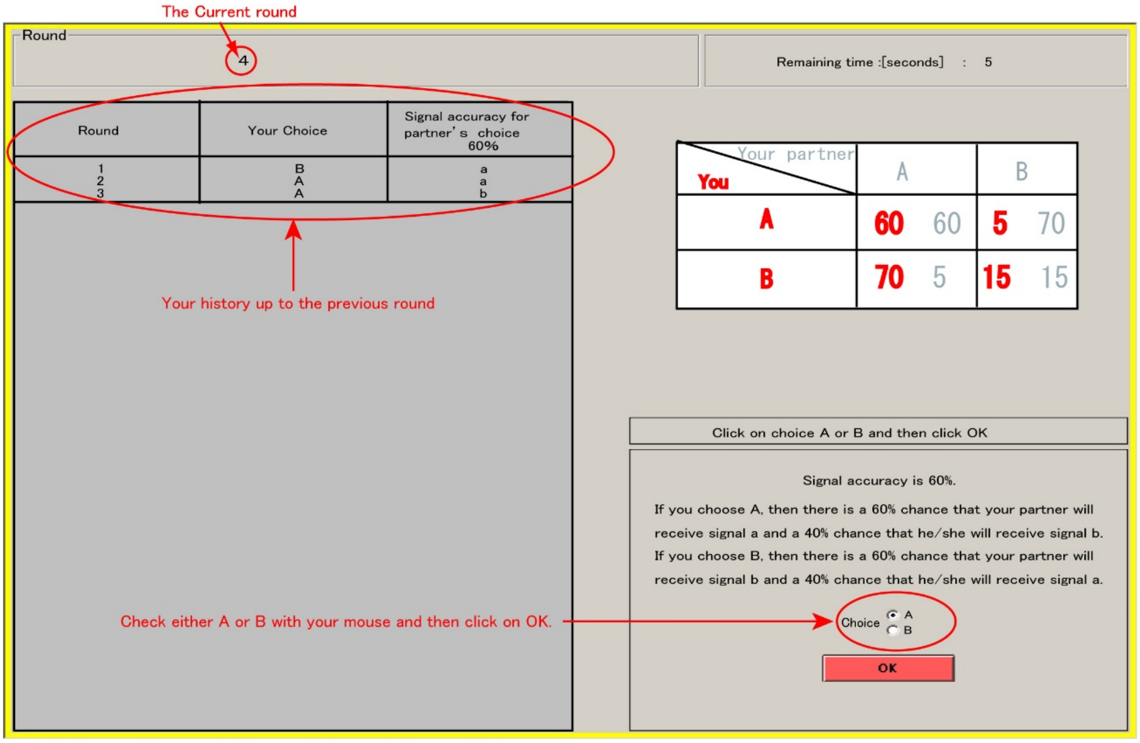
Screen 7: The Results Screen



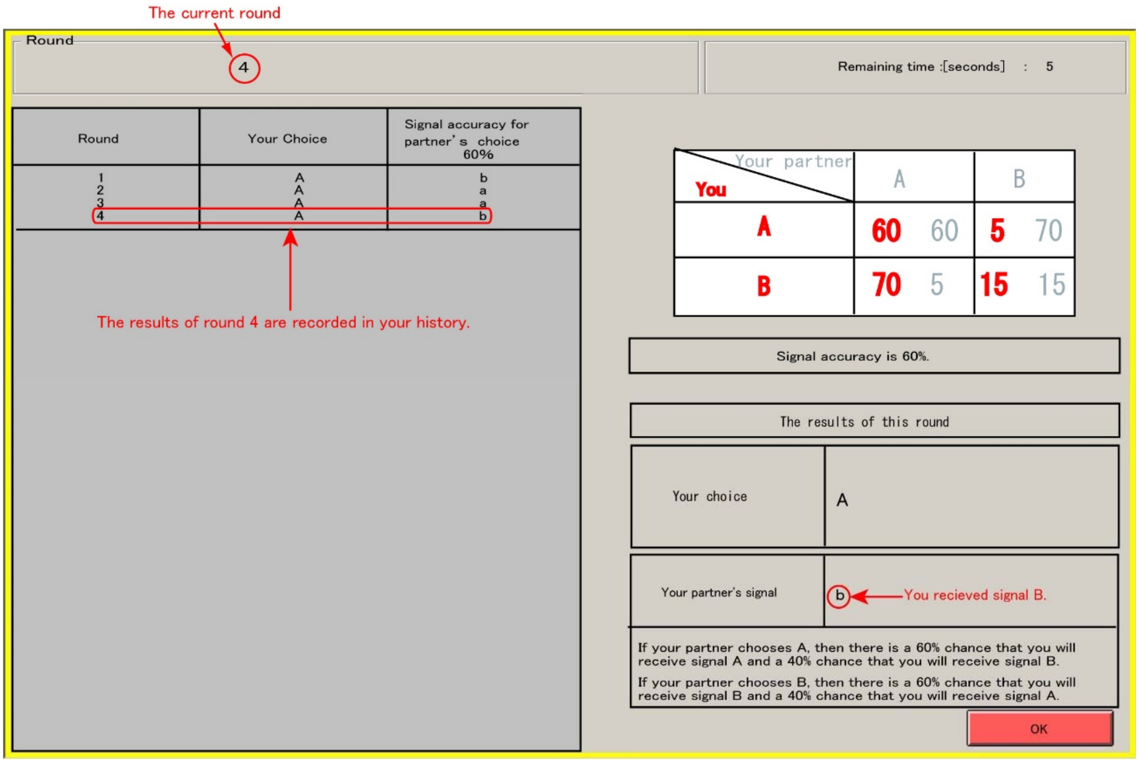
Screen 8: Your Selection Screen



Screen 9: The Selection Screen of Your Partner



Screen 10: Your Results Screen



Screen 11: The Results Screen of Your Partner