Learning about a Reciprocating Opponent in an Iterated Prisoner's Dilemma

A Dissertation Presented

by

Pei-Pei Liu

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Psychology

(Experimental Psychology)

Stony Brook University

December 2014

Stony Brook University
The Graduate School

Pei-Pei Liu

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Christian Luhmann – Dissertation Advisor
Associate Professor, Department of Psychology

Suparna Rajaram – Chairperson of Defense
Professor, Department of Psychology

Antonio Freitas – Committee Member
Associate Professor, Department of Psychology

Reuben Kline, Outside Member
Assistant Professor, Department of Political Science

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Learning about a Reciprocating Opponent in an Iterated Prisoner's Dilemma

by

Pei-Pei Liu

Doctor of Philosophy

in

Psychology

(Experimental Psychology)

Stony Brook University

2014

Research has shown that reciprocation increases individuals' willingness to cooperate. This study investigates how individuals learn to cooperate with reciprocating opponents. To do so, we evaluated individuals' expectations about the behavior of their opponents during an iterated Prisoner's Dilemma (PD). In four experiments, participants played with a Tit-For-Tat (TFT) algorithm that occasionally failed to reciprocate. In Experiment 1, we first established whether individuals actually develop expectations about their opponents by utilizing a concurrent task. Our results indicate that when the opponents did not reciprocate, participants engaged in greater cognitive processing and were slower to respond to the concurrent task. Experiment 2 examined whether delayed reciprocation affects expectations about reciprocation using similar methodology. Our results indicate that expectations were weaker when reciprocation was delayed. In Experiment 3, we investigated two possible paths through which people may learn to cooperate with TFT. Specifically, we investigated whether the expectations people develop concern their own payoffs or the behavior of their opponents. Our results indicate that participants' expectations concern both their own payoffs and opponents' behavior. In Experiment 4, we sought for convergent evidence and a finer temporal resolution by employing pupillometry. Our results indicate that participants exhibited greater pupil sizes when expectations about reciprocation were violated.

Table of Contents

- Introduction 1
  - Learning to Cooperate in PD 2
  - The Delay Associated with Reciprocation 4
  - Two Potential Learning Paths 6
  - Measuring Expectations 7
  - The Current Project 9
- Experiment 1 9
  - Methods 10
  - Results 13
  - Discussion 13
- Experiment 2 14
  - Methods 14
  - Results 15
  - Discussion 16
- Experiment 3 16
  - Methods 17
  - Results 19
  - Discussion 20
- Experiment 4: Pupillometry........20
  - Methods 22
  - Results 23
  - Discussion 24
- General Discussion 25
- References 29

List of Figures

- Figure 1 …..…………………………………………………….............……31
- Figure 2 ……..…………………………………………………………...32
- Figure 3 ……..…………………………………………………………...33
- Figure 4 ……..…………………………………………………………...34
- Figure 5 ……..…………………………………………………………...35
- Figure 6 ……..…………………………………………………………...36
- Figure 7 ……..…………………………………………………………...37
- Figure 8 ……..…………………………………………………………...38
- Figure 9 ……..…………………………………………………………...39

Introduction

Individuals often encounter situations in which they have to choose between behaving selfishly and cooperating for the mutual benefit of some larger group. For example, one might selfishly litter, but everyone would be better off if she recycled. These decisions can be difficult because the ultimate consequences are uncertain; the consequences are jointly determined by multiple individuals each faced with the same choices. In these situations, any given individual achieves her own best outcome when she behaves selfishly but others cooperate. However, when everyone in the situation behaves selfishly, these individuals as a group are worse off. Because pure self-interest is unlikely to yield collectively optimal outcomes, it is important to understand the factors that affect such decisions.

Prisoner's Dilemma (PD hereafter) is one of the most studied models of cooperation. A standard PD involves two players who can each choose to cooperate or to defect (see Figure 1). Each player achieves her highest payoff when she defects and the other player cooperates, and achieves her lowest payoff when she cooperates and the other player defects. When both players choose to cooperate, payoffs are moderately high for each player. When both players choose to defect, payoffs are moderately low for each player. Close inspection of Figure 1 suggests that the payoffs for defection are always higher than the payoffs for cooperation, regardless of what the other player chooses. When both players adopt this perspective, the most likely outcome is mutual defection. However, as mentioned earlier, the payoff for mutual defection is lower than the payoff for mutual cooperation. When two players play the game repeatedly, the difference between mutual defection and mutual cooperation adds up quickly, and thus decision makers may view cooperation as a better option than defection.

Given that the iterated version of the PD would seem to provide incentives for mutual cooperation, one might be surprised to find that past research has found that mutual defection actually occurs far more often in iterated PDs than mutual cooperation (Duffy & Smith, 2012; Rachlin, Brown, & Baker, 2000; Rapoport & Chammah, 1965; Roth, 1995; Silverstein, Cross, Brown, & Rachlin, 1998). Cooperation rates as low as 20% have been reported (e.g., Duffy & Smith, 2012; Rachlin et al, 2000; Silverstein et al., 1998). However, research also shows that individuals can learn to cooperate when repeatedly playing against an opponent who reciprocates (Bakers & Rachlin, 2001; Silverstein et al., 1998; Fantino, Gaitan, Meyer, & Stolarz-Fantino, 2006; Kuhlman & Marshello, 1975; Stephens, McLinn, & Stevens, 2002, 2006; Yi, Johnson, & Bickel, 2005).

Learning to Cooperate in PD

Reciprocation is best illustrated by a strategy called Tit-for-Tat (TFT; Axelrod, 1984). A standard TFT algorithm cooperates on the first round in an iterated PD and then chooses what the opponent has chosen on the previous round. As can be seen in Figure 1, an individual is always better off when her opponent cooperates than when her opponent defects. Therefore, TFT rewards cooperation by cooperating on the next round and punishes defection by defecting on the next round. In this way, TFT can be thought of as an algorithm which attempts to teach its opponent to cooperate in an iterated PD. Looking at it in another way, when playing against a TFT player, switching back and forth between cooperation and defection yields an average payoff that is the average of the highest payoff (e.g., 6 points in Figure 1) and the lowest payoff (e.g., 1 point in Figure 1). In contrast, making the same choice repeatedly will cause the TFT player to also make the same choice repeatedly, resulting in either repeated mutual cooperation (yielding 5 points per round) or repeated mutual defection (yielding 2 points per round). Therefore, to maximize one's overall payoff against a TFT player, one must cooperate repeatedly. In this way, TFT algorithm provides incentives for cooperation and disincentives for defection.

Consistent with the rationale behind TFT, research has shown that individuals do cooperate more when playing against TFT than when playing against other, less reciprocative strategies (e.g., Baker & Rachlin, 2001; Fantino et al., 2006; Kuhlman & Marshello, 1975; Silverstein et al., 1998). For example, Baker and Rachlin (2001) had participants play an iterated PD against opponents that reciprocated on 50%, 75%, or 100% of rounds (100% being equivalent to TFT), choosing randomly otherwise. Baker and Rachlin found that participants cooperated more (approximately 75% cooperation) when opponents reciprocated 100% of the time than when opponents reciprocated less consistently (approximately 30% cooperation). Other studies have compared TFT with other strategies, such as those that either always defect or always cooperate (Kuhlman & Marshello, 1975; Silverstein et al., 1998). These studies have found that cooperation levels were higher against TFT than against these other strategies. These findings suggest that people are sensitive to reciprocation and are more willing to cooperate when their opponents consistently reciprocate.

Research has repeatedly shown that TFT indeed boosts cooperation, suggesting that individuals learn what TFT attempts to teach them (i.e., to cooperate). Although there has been discussions regarding why TFT might elicit cooperation, little discussion has taken the perspective of the learner; that is, the opponents playing against TFT. When playing against a TFT opponent, what exactly are people learning? What are the psychological processes that allow individuals to actually learn to cooperate when interacting with TFT opponents? As will be pointed out later, there are multiple paths through which individuals could learn to cooperate with a reciprocating opponent. The current study attempts to answer these questions.

The Delay Associated with Reciprocation

One potential difficulty in learning to cooperate with a reciprocating opponent is that reciprocation is necessarily delayed. That is, an opponent's behavior can only be reciprocated after the behavior has occurred. Therefore, reciprocation of current behavior must wait at least until the next interaction. Such delays may hurt the effectiveness of reciprocation because delayed incentives are less effective behavioral reinforcers (Kamin, 1954; Gallistel & Gibbon, 2000; Sutton & Barto, 1998; Weinberg, Luhmann, Bress, & Hajcak, 2013; for a review, see Renner, 1964). Consistent with this idea, studies have shown that the longer the delay associated with reciprocation is, the less effective it is in encouraging cooperation (Baker & Rachlin, 2002; Fantino et al., 2006; Komorita, Hilty, & Parks, 1991; Locey & Rachlin, 2011; Liu & Luhmann, 2014a; Parks & Rumble, 2001; Stephens et al., 2002). However, most of these studies adopted either a non-standard version of TFT (Komorita et al., 1991; Parks & Rumble, 2001) or a non-standard version of PD (Baker & Rachlin, 2002; Fantino et al., Locey & Rachlin, 2011), or investigated the behavior of nonhuman animals (Baker & Rachlin, 2002; Stephens et al., 2002). For example, Komorita et al (1991) used a non-standard TFT which delayed reciprocation by two rounds instead of by only one round. However, interpretation of their results was complicated by the fact that the two-round delay was only implemented either when participants switched from cooperation to defection or when they switched from defection to cooperation. In contrast, Locey and Rachlin (2011) used a variant of PD in which each player knew the other player's choice before she made her own choice. It is unclear how the effects of delayed reciprocity found using these non-standard procedures generalize to more standard procedures in human participants.

To address this interpretational ambiguity, we (Liu & Luhmann, 2014a) have recently examined the effect of delayed reciprocation in a standard PD, with human participants playing against a standard TFT algorithm. We manipulated the delay associated with reciprocation by manipulating the temporal intervals between rounds (i.e., inter-trial interval, ITI) in iterated PDs against TFT. We found that the overall cooperation level was lower when the ITI was long (6 seconds) than when the ITI was short (.5 seconds). Moreover, participants in the short ITI condition cooperated more as they repeatedly interacted with their TFT opponent. In contrast, participants in the long ITI condition showed no such changes. These findings suggest that learning to cooperate against a reciprocating opponent is undermined when reciprocation is delayed.

These findings leave several important questions unanswered. For example, why does delaying reciprocation reduce cooperation? What cognitive processes are affected by such delay? To address this question, Liu and Luhmann (2014a) investigated the importance of memory in learning of cooperation. Several aspects of memory have been discussed (Stevens & Hauser, 2004; Stevens, Cushman, & Hauser, 2005) as being important in detecting reciprocation, which requires keeping track of the history of earlier interactions with an opponent. Specifically, in order to determine that an opponent is reciprocating, one must detect the relationship between one's own past choices and the opponent' current choices. Therefore, memory of one's own previous choices should be crucial in understanding reciprocation, which in turn, should affect one's subsequent decisions regarding cooperation. To reduce such memory demands we provided a reminder of the choice made by the participant on the previous round. We found that participants provided with such reminders cooperated more than those who were not provided with the reminder. This finding suggests that memory for one's earlier choices is important when learning to cooperate against a reciprocating opponent.

Liu and Luhmann noted that there are at least two possible ways in which memory of one's own previous choice can facilitate learning of cooperation. First, in order to detect reciprocation on the part of one's opponent, one has to see the relationship between the opponent's current choice and one's own choice on the previous round. When reciprocation is delayed more (e.g., when the ITI was long), it may be more likely that one will fail to remember one's own prior choice and thus fail to recognize that the opponent is reciprocating. Second, when playing against TFT, one's current cooperation is associated with future cooperation on the part of the opponent, which is in turn associated with greater payoffs. Likewise, one's current defection is associated with future defection on the part of the opponent, which is associated with smaller payoffs. Thus, the more delayed reciprocation is, the more difficult it is to associate one's current behavior with the future payoffs that come with reciprocation, and thus reciprocation may be less effective in encouraging cooperation.

Two Potential Learning Paths

The above explanations for the memory effect observed in Liu and Luhmann (2014a) point out two possible ways by which people might learn to cooperate against a TFT opponent. First, players may recognize that their opponents are reciprocating, and hence realize that it is in their best interest to cooperate with such opponents. On the other hand, recognition of reciprocation per se is not necessary to explain the above findings. Players may instead simply learn that cooperation is more valuable over the long term. As discussed earlier, cooperation with a TFT opponent predicts greater payoffs in the long run than does defection. Specifically, when one cooperates with TFT, her payoffs on the next round will be higher (e.g., 5 or 6 in Figure 1) than when she defects (e.g., 1 or 2). The difference between the higher (i.e., 5.5) and lower (i.e., 1.5) future payoffs on the next round is greater than the difference between the immediate payoffs associated with defection (i.e., 4) and cooperation (i.e., 3) on a current round. That is, defection leads to a small gain immediately but a greater loss later. Therefore, one may view cooperation as ultimately more rewarding and would thus be more likely to cooperate when interacting with a reciprocating opponent. In the current study, we examine these two alternative explanations.

To investigate how individuals learn to cooperate, the current project focuses on whether people develop expectations as they interact with TFT opponents. Specifically, if one recognizes the reciprocating nature of her TFT opponent, she will learn the association between her own choices and her opponent's subsequent choices. In this case, she develops expectations about the opponent's behavior on each round according to the principle of reciprocation. When the TFT opponent fails to reciprocate on a specific trial, the player's expectation of reciprocation is violated. On the other hand, if one simply associates cooperation with greater subsequent payoffs, she will develop expectations about the payoffs she will earn. In that case, if her TFT opponent fails to reciprocate on a specific trial, she should be surprised because she does not receive the expected payoffs.

Measuring Expectations

Violation of expectation, or prediction error, is a critical component in many theories of learning (e.g., Catena et al., 1998; Hogarth & Einhorn, 1992; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). These theories suggest that learners make predictions about the future events based on prior experience. Learning occurs when such predictions are violated; when learners are surprised. Such violations suggest that the learner's beliefs about the world are flawed and thus in need of correction. When events occur as predicted, learners are not surprised and no learning takes place.

Research in contingency learning has shown that one consequence of expectation violations is that learners engage greater cognitive processing. For example, Liu and

Luhmann (2014b) presented participants with a series of trials, the majority of which presented evidence consistent with a strong, positive relationship between two events (e.g., the occurrence of event A was reliably followed by the occurrence of event B). Intermixed in the trial sequence were a minority of trials that were consistent with the opposite relationship (e.g., the occurrence of event A was followed by the absence of event B). The design of the sequence was such that participants were predicted to develop expectations based on the majority of evidence (e.g., that event B follow event A) and that such expectations would be violated on a minority of trials. As participants completed the learning task, auditory tones were occasionally played and participants were instructed to respond based on the pitch of the presented tone. Liu and Luhmann found that responses to the tones were slower on the minority trials, suggesting that observations inconsistent with previous information elicited greater cognitive processing than observations consistent with previous information. In addition, they found that the magnitude of the response time effect was correlated with the degree to which participants learned the relationship between events.

The tone-discrimination task also allows us to monitor how learning processes unfold over time. For example, we have utilized this measure in a blocking paradigm (Liu & Luhmann, 2013) as well as in paradigms that manipulate presentation order (Liu & Luhmann, 2014b). This measure was able to provide more direct evidence about cognitive processes that unfold on a trial-by-trial basis. For example, the findings from Liu and Luhmann (2013) suggest that there is a transient increase in cognitively demanding processing during a blocking paradigm that is strongly related to learning itself.

The Current Project

In the current study, we utilized the tone-discrimination task to monitor players' expectations about the behavior of their TFT opponents during an iterated PDs. In

Experiment 1, we investigated whether players actually develop expectations during an iterated PD against TFT opponents. If players indeed develop expectations about their opponents' reciprocation, when their TFT opponents occasionally do not reciprocate, these expectations would be violated, which should elicit greater processing and slow down responses to the tone probes. In Experiment 2, we examined whether delayed reciprocation, which has been shown to undermine cooperation, affects expectations about reciprocation. In Experiment 3, we directly investigated the possible paths through which people learn to cooperate with a reciprocating opponent. To do that, we investigated whether the expectations people develop throughout the sequence of interactions concern their own payoffs or the behavior of their opponents. In Experiment 4, we sought convergent evidence and potentially more sensitive measurements by employing pupillometry.

Experiment 1

In Experiment 1, we first examined whether players develop expectations during an iterated PD when playing against a TFT opponent. To test for expectations, the TFT algorithm did not reciprocate on a small number of rounds (i.e., the non-TFT rounds) and instead did the opposite of what the standard TFT opponent would do. Following Liu and Luhmann (2013, 2014b), auditory tones were played on a small number of TFT rounds and a small number of non-TFT rounds. If participants develop expectations about their opponent's reciprocation, responses to the tones should be slower on non-TFT trials as a consequence of players' expectations being violated.

Methods

Participants

Thirty-six Stony Brook University undergraduates participated in this experiment for partial course credit. Following Locey and Rachlin (2012) and our previous work (Liu & Luhmann, 2014a) we adopted a policy of excluding any participant who failed to sample each choice (i.e., cooperation and defection) at least four times. Because the experiment was designed to study how players learn about reciprocative behavior, individuals who did not sample both choices never had an opportunity to engage in such learning. Four participants were excluded from analyses based on this policy.

Materials & Design

Participants played the Prisoner's Dilemma game with the payoffs depicted in Figure 1. In order to avoid the influence of overt social norms (e.g., positive connotations of "cooperation"), green and blue cards were used to represent cooperation and defection respectively. The payoff information was presented in the format shown in Figure 2, illustrating how the two players' choices jointly determined their earnings.

Unknown to participants, their opponent was actually a computer algorithm. Each participant completed a total of 100 rounds of the game. On 90 of these rounds, the algorithm played standard TFT (i.e., the TFT rounds). On these TFT rounds, the algorithm selected what the participant chose on the previous round (TFT will always cooperate on the first round). On the other 10 rounds, the algorithm did the opposite of what TFT should do; that is, it selected the opposite of what the participant chose on the previous round. The first 20 rounds of a sequence of PD consisted entirely of TFT rounds so that participants had the opportunity to develop strong expectations based on pure reciprocation. The non-TFT rounds were intermixed with the TFT rounds in the last 80 rounds of the sequence. The order of the two kinds of rounds was pseudo-random to ensure that the non-TFT rounds were spread out over the sequence. Specifically, the last 80 trials were divided into 10 eight-round blocks, with one non-TFT round and seven TFT rounds in each block. The order of rounds within each block was randomized.

The tones delivered throughout the sequence of iterated PD consisted of three different frequencies (just as in Liu & Luhmann 2013, 2014b). The high tone had a frequency of 3520Hz, the medium tone had a frequency of 880Hz, and the low tone had a frequency of 220Hz. Each tone lasted for 50 ms. A tone was played during two of the first 20 rounds of the sequence. During the last 80 rounds, a tone was played during eight TFT rounds and six non-TFT rounds. The frequencies of the tones were randomly determined on each trial.

Procedure

Participants were run in pairs. Each pair of participants was given task instructions together and was led to believe that they would play a game with each other. Participants were told that their goal was to earn as many points as they could. Participants were instructed to press the arrow keys on the keyboard to indicate their choice. They had to press the left arrow key to select the green card (i.e., cooperation) and the right arrow key to select the blue card (i.e., defection). Participants were also told that whenever they heard a tone, they had to press one of three keys (the "1", "2", and "3" keys) according to the pitch of the tone.

Before starting, participants were first familiarized with the tone probes, following the procedure used in Liu and Luhmann (2013, 2014b). A sequence of the three different tones were presented randomly and participants were required to simply identify whether the presented tone was high, medium, or low using the three keys on the computer keyboard. Feedback was presented after each response and participants continued until they were able to discriminate the three tones.

So as to enhance the impression of playing with an actual human opponent in the iterated PD, the screen displayed two boxes representing the choices from each player (see Figure 2). A question mark was presented inside each box at the beginning of the trial. When the participant made her selection, the corresponding question mark changed to a check mark, indicating that the participant had made a choice. The question mark in the opponent's box changed to a check mark 2.5 seconds after the onset of the round (i.e., as though the opponent had taken 2.5 seconds to make her choice). Once both the participant and the computer made their choices, the payoff information was removed. After 1.5 seconds, the outcome of that round was revealed. The boxes displayed the selected choices and the corresponding payoffs for both the participant and the participant's opponent. After 2 seconds, the screen was cleared. The next round of the game started after an additional 2 seconds.

On probed rounds, the tone was presented one second after the outcome was displayed. Participants then had up to three seconds to respond to the tone and the outcome of that round remained on the screen until participants responded. If no response was made within three seconds, the round would end and the task moved on to the next round.

Results

To determine whether participants develop expectations about reciprocation in an iterated PD, we compared response times (RTs) to tones played on TFT rounds and to tones played on non-TFT rounds. Because the first 20 trials consisted only of TFT rounds, RTs to the tones delivered during this period were not used in the analyses. As shown in Figure 3, participants were slower to respond to tones played on non-TFT rounds (M = 1.42, SD = 0.27) than to tones played on TFT-rounds (M = 1.32, SD = 0.19, $t(31) = 3.09$, $p < .01$), suggesting that participants were engaged in more intense cognitive processing on non-TFT rounds than on TFT-rounds. This result suggests that participants developed expectations about reciprocation during the iterated PD against the TFT opponent and such expectations were violated on non-TFT rounds.

Overall, participants cooperated on 43% of the rounds (SD = 0.17). To examine how cooperation rates changed over time, we divided the 100-round sequence into five blocks, each consisting of 20 rounds. As shown in Figure 4, cooperation rates decreased over time ($F(4, 74.75) = 6.36$, $p = .002$).

Discussion

Experiment 1 found that participants were slower to respond to tones when the otherwise reciprocative opponent failed to reciprocate, suggesting that participants engaged in deeper processing on these non-reciprocating rounds. Such a finding is consistent with the idea that participants develop expectations about reciprocation over the course of repeated interactions with a reciprocating opponent. This finding also establishes that the concurrent tone-discrimination task can be used to track expectations about reciprocation in an iterated PD. The decrease in cooperation rates observed over the sequence of interactions may be attributed to the insertion of non-TFT rounds. As demonstrated in previous research, although reciprocation encourages cooperation, probabilistic reciprocation is not particularly effective (e.g., Baker & Rachlin, 2001).

Experiment 2

Prior research has demonstrated that cooperation against a reciprocative opponent is undermined when reciprocation is delayed (e.g., Baker & Rachlin, 2002; Liu & Luhmann, 2014a). Delaying reciprocation may hinder cooperation either because it makes it difficult for individuals to recognize reciprocation or because it makes it more difficult to associate current cooperation with higher future payoffs. However, in either case, delaying reciprocation should weaken the expectations that develop as they interact with a reciprocative opponent. Experiment 2 attempted to examine this prediction by manipulating the delay associated with reciprocation and investigating whether such a manipulations altered participants' expectations.

Methods

Participants

Eighty Stony Brook University undergraduates participated in this experiment for partial course credit. Participants were randomly assigned to either the Long-ITI condition or the Short-ITI condition. Nine participants in the Long-ITI condition and seven participants in the Short-ITI condition failed to sample each choice at least four times and were excluded from subsequent analyses, leaving 32 participants in each condition.

Design and procedure

The design and procedure of the iterated PD task were identical to those used in Experiment 1 except for the following difference: in the Short-ITI condition, rounds were separated by 0.5 seconds whereas in the Long-ITI condition, rounds were separated by six seconds. These delays were identical to those used in Liu and Luhmann (2014a).

Results

To examine whether the delay associated with reciprocation influenced the strength of expectations about reciprocation, we compared the magnitudes of the RT effects in the Long-ITI and Short-ITI conditions (see Figure 4). We conducted a 2 (round type: TFT vs. non-TFT) by 2 (ITI: Long vs. Short) mixed ANOVA on RTs to the tone probes, with repeated measures on the former factor. The results indicated a main effect of ITI ($F(1, 62) = 9.33$, $p = .003$) and a marginal main effect of types ($F(1, 62) = 3.30$, $p = .07$). The interaction was also marginally significant ($F(1, 62) = 2.40$, $p = .13$). Further analyses suggest that in the Short-ITI condition, participants were slower to respond to tones played on non-TFT rounds (M = 1.30, SD = 0.28) than to tones played on TFT-rounds (M = 1.23, SD = 0.22, $t(31) = 2.19$, $p = .04$). In contrast, in the Long-ITI condition, RTs to tones played on non-TFT rounds (M = 1.45, SD = 0.25) and TFT-rounds (M = 1.44, SD = 0.26) were not different ($t(31) < 1$, $p > .8$). Taken together, these results suggest weaker expectations about reciprocation in the Long-ITI condition than in the Short-ITI condition.

We next examined whether ITI influenced cooperation rates. We divided the interaction sequence into five blocks, each consisting of 20 rounds. As shown in Table 5, cooperation rates decreased in both the Long-ITI and Short-ITI conditions. This observation is confirmed by a 2 (ITI: Long vs. Short) by 5 (block) mixed ANOVA, with repeated measure on the latter factor. The results indicated a main effect of block ($F(4, 248) = 8.35$, $p < .001$). The main effect of ITI ($F(1, 62) = 1.38$, $p = .24$) and the interaction between the two factors ($F(4, 248) = .53$, $p = .72$) were not significant.

Discussion

Experiment 2 found that participants were slower to respond to tones played on non-TFT rounds than on TFT rounds, but only when the ITI was relatively short. This finding suggests that participants were more sensitive to violations of reciprocation when reciprocation was associated with shorter delays. That is, participants' expectations about reciprocation appear to be attenuated when reciprocation is delayed. This finding extends on previous findings that delaying reciprocation impedes participants' tendency to cooperate (e.g., Baker & Rachlin, 2002; Liu & Luhmann, 2014a) and suggests that delays hurt cooperation because participants were not learning about reciprocation as well.

Consistent with Experiment 1, participants in both conditions of Experiment 2 exhibited decreases in cooperation over time. Unlike in Liu and Luhmann (2014a), there was no influence of ITI on cooperation rates. This is likely due to the inclusion of non-TFT rounds, which may have acted to reduce cooperation (e.g., Baker & Rachlin, 2001). Moreover, such influence of the non-reciprocating rounds may have overridden the influence of ITI and resulted in decreased cooperation regardless of ITIs.

Experiment 3

In Experiment 3, we tested the two potential paths by which individuals can learn to cooperate with a reciprocating opponent. As discussed earlier, it is possible that participants become more cooperative because they recognize reciprocation in their opponents. On the other hand, participants may become more cooperative because they directly learn that cooperation is more rewarding in the long run. In the first case, participants' learning focuses on the association between their own current choices and opponents' future choices (i.e., a choice-choice association). Therefore, participants should develop expectations that primarily concern opponents' choices. In the second case, participants' learning focuses on the association between their own current choice and their own future payoffs (i.e., a choice-payoff association). As a result, participants should develop expectations that primarily concern their own payoffs.

To examine these alternatives, we used the tone probe paradigm to measure participants' expectations about specific aspects of the iterated PD. Specifically, we presented tone probes in the presence of either the choices selected by each player (Choice-Focus condition) or the payoffs (Payoff-Focus condition). If participants cooperate because of the choice-choice association, non-TFT rounds should violate expectations in the Choice-Focus condition but not in the Payoff-Focus condition. On the other hand, if participants cooperate because of the choice-payoff association, non-TFT rounds should violate expectation in the Payoff-Focus condition but not in the Choice-Focus condition.

Methods

Participants

Seventy Stony Brook University undergraduates participated in this experiment for partial course credit. Participants were randomly assigned to either the Choice-Focus condition or the Payoff-Focus condition. Four participants failed to sample each choice at least four times and were excluded from subsequent analyses, leaving 34 participants in the Payoff-Focus condition and 32 participants in the Choice-Focus condition.

Design & Procedure

The design and procedure of the iterated PD task were identical to those used in Experiment 1 except for the following differences. First, when displaying the outcome, 14 TFT-rounds and six non-TFT rounds presented outcomes such that attention should be focused on either the selected choices or on the payoffs (depending on the condition). These rounds are referred to as the focus rounds and Figure 6 shows the procedure of displaying the outcome on these rounds in each condition. On focus rounds in the Choice-Focus condition, the cards selected by each player were presented but the corresponding payoffs were not. After two seconds, the payoffs were also presented along with the selected cards for an additional two seconds. Similarly, on the focus rounds in the Payoff-Focus condition, payoffs for both players were presented first without showing the selected cards. After two seconds, the selected cards were displayed along with the payoffs for an additional two seconds. Therefore, there were 20 focus rounds and 80 standard rounds (which consist of the same outcome displays as in Experiment 1). The two conditions differ only in the outcome displays on the focus rounds. Second, tones were played on 14 focus rounds (eight TFT and six non-TFT rounds) and four standard rounds. Therefore, both standard and focus rounds were probed, and not all focus rounds were probed. In this way, participants could not anticipate tone probes based on the nature of the round. Third, when focus rounds were probed, the tone was delivered one second after the display of the selected cards in the Choice-Focus condition and one second after the display of the payoffs in the Payoff-Focus condition. Participants had up to 3 seconds to respond. If a response to the tone was made within 3 seconds, the full outcome information (i.e., both the selected cards and the payoffs) was shown immediately after the response. Otherwise the task moved on to show full outcome information at the conclusion of the 3-second response window.

Results

To examine the relative importance of the two learning paths, we compared the RT effects in the Choice-Focus and Payoff-Focus conditions. For the purpose of this analysis, only RTs to tones played on the focus rounds were used. A 2 (round type: TFT vs. non-TFT) by 2 (condition: Choice-Focus vs. Payoff-Focus) mixed ANOVA was conducted on RTs to the tones, with repeated measures on round type. The analysis indicated a main effect of round type ($F(1, 64) = 20.59$, $p < .001$). The main effect of condition and the interaction between the two factors were not significant (F's < 1). Further analyses indicated that responses to the tones played on non-TFT rounds were slower than response on TFT rounds in the Choice-Focus condition (M = 1.45 vs. 1.34, $t(31) = 3.89$, $p < .01$) as well as in the Payoff-Focus condition (M = 1.39 vs. 1.30, $t(33) = 2.68$, $p < .001$; Figure 7).

To examine whether initially revealing only partial outcome information on the focus rounds influenced responses to tones, we compared RTs to tones on focus rounds and RTs to tones on standard rounds (using only TFT rounds). A 2 (trial: focus vs. standard) by 2 (condition: Choice-Focus vs. Payoff-Focus) mixed ANOVA was conducted on RTs to the tones, with repeated measures on trial. The analysis indicated that responses were slower on standard rounds (M = 1.37, SD = 0.28) than on focus rounds (M = 1.32, SD = 0.28, $F(1, 64) = 4.22$, $p < .05$). The main effect of condition and the interaction were not significant (p's > .5). These results suggest that participants engaged in greater cognitive processing when presented with complete information than when presented with only partial information about outcome.

Next we examined cooperation over time. Again, the sequence of interactions was divided into five blocks, each consisting of 20 rounds. Figure 8 shows an overall decrease in cooperation in both conditions, with no obvious difference between the two conditions. A 5 (time blocks) by 2 (condition: Choice-Focus vs. Payoff-Focus) mixed ANOVA was conducted on cooperation rates, with repeated measures on round type. The analysis revealed a main effect of time blocks ($F(4, 256) = 25.47$, $p < .001$). The main effect of condition and the interaction were not significant (F's < 1). These results indicate that cooperation decreased over time regardless of whether one's own payoff or opponents' choice was emphasized on focus rounds.

Discussion

Experiment 3 found that participants were slower to respond to tones on non-TFT rounds than on TFT rounds, regardless of whether they were viewing information regarding their own payoffs or information regarding their opponents' choice. This finding suggests that participants develop expectations regarding both their own payoffs and their opponents' behavior. When the opponent failed to reciprocate, participants found both their own payoffs and the opponents' behavior surprising. Furthermore, we found no difference in the magnitude of the RT effects in Choice-Focus and Payoff-Focus conditions, suggesting that participants' expectations about their own payoffs were just as strong as their expectations about opponents' behavior. Such a finding suggests that participants did not exclusively rely on one learning path or the other. Instead, it appears that participants learned the choice-payoff association and the choice-choice association equally well.

Experiment 4: Pupillometry

The first three experiments investigated how individuals learn to cooperate by using tone probes to measure people's expectations. Experiment 4 sought to investigate this same question using a physiological measure, pupil diameter specifically. Physiological measures such as pupillometry are of interest because they are not generally under voluntary control. In addition, such measures require no voluntary responses from participants and are therefore less likely to interfere with the learning processes of interest. Moreover, the high temporal resolution of pupillometry provides insight to the temporal dynamics regarding the computation of expectation violations. For example, depending on how automatic the computation of expectation violation is, physiological effects may emerge almost immediately after the outcome of a round is revealed. In contrast, if violations require more complex processing (e.g., several steps of inference), effects may be slower to emerge. Thus, pupillometry can describe the learning processes in the iterated PD to a more detailed extent than the tone probe methodology employed in Experiments 1-3.

Research has shown that pupil dilations reflect processing efforts in a variety of contexts (for reviews, see Beatty & Lucero-Wagoner, 2003, and Laeng, Sirois, & Gredeback, 2012). Pupil diameter has been demonstrated to increase when difficulty increases in a wide range of cognitive tasks such as signal discrimination, target detection, memory span, and language processing, as well as when viewing emotional pictures. Given that our own research has suggested that expectation violations during learning increase cognitive processing (Liu & Luhmann, 2013, 2014b), it is reasonable to expect that pupil diameter will also reflect violated expectations. Indeed, Satterthwaite, Green, Myerson, Parker, Ramaratnam, and Buckner (2007) found that pupil diameter increased when participants encountered an unexpected loss in a decision-making task, suggesting that pupil dilations may reflect violation of expectations. Therefore, in Experiment 4, we utilized pupil dilations to measure violations of expectations about reciprocation.

In Experiment 4, we examined whether pupil dilations indeed capture violations of expectation in an iterated PD. Instead of playing tone probes, we utilized changes in pupil size to track expectations. Experiment 4 employed the same design as Experiment 1. If participants develop expectations as they interact with a reciprocating opponent, pupil sizes should be greater on non-TFT rounds than on TFT rounds.

Methods

Participants

Thirty-five Stony Brook University undergraduates participated in this experiment for partial course credit. All participants had normal or corrected-to-normal vision. Three participants failed to sample each choice at least four times and were excluded from subsequent analyses.

Design & Procedure

The design and procedure was identical to those in Experiment 1 except for the following changes. First, no tone probes were played. Second, a flicker fusion task was used to match the luminance in green and blue for each participant. During this task, a circular stimulus (in either green or blue) was repeatedly presented on a gray background for a brief interval (with each presentation lasting 50 milliseconds). Participants pressed two keys to increase and decrease the luminance of the stimulus until the stimulus stopped flickering. That stopping point represented the point at which the luminance of the stimulus was as subjectively equivalent to the luminance of the background as the colors would allow. Every participant completed this procedure with the green and the blue stimulus respectively. The two colors selected in this task were then used as the color of the two cards in the iterated PD for that participant. Third, a fixation cross was presented in the center of the screen beginning one second before each round started and remained throughout the whole round. Third, on each round after both players made a choice but before displaying the outcome, the check marks and the payoff matrix were removed from the screen, leaving only the fixation cross and the "You" and "Opponent" labels on the screen. This remained for 2 seconds, after which the outcome of the round was displayed for 2 seconds. Pupil diameter was measured during the one second preceding the outcome (i.e., baseline period) as well as during outcome display itself.

Stimuli &Apparatus

The stimuli were identical to those used in Experiment 1 except for the following changes. First, stimuli were presented over a neutral grey background instead of a white background. The colors used to represent the two choices for each participant were chosen using the flicker fusion task described above. Second, a SR Research EyeLink 1000 eye tracking system was used to measure participants' pupil sizes. The system has a sampling rate of 1000 Hz and a spatial resolution of 0.25° to 0.5°. Stimuli were displayed on a flatscreen CRT monitor. Head position and viewing distance were fixed with a chin rest approximately 69 cm from the monitor.

Results

Before performing statistical analyses, we first subtracted the average pupil diameter during the baseline period from the diameter during the outcome display. This was done separately for each round and for each participant. This procedure ensured that our measurements represent a task-related change in pupil diameter and not trial-to-trial fluctuations. We then used the baselined pupil diameters to examine differences between TFT and non-TFT rounds.

We compared the pupil diameter during outcome display on TFT rounds with pupil diameter on non-TFT focus rounds. The two-second period of outcome display was divided into ten 200-msec time blocks. A 2 (round type: TFT vs. non-TFT) by 10 (blocks) repeated ANOVA was conducted on pupil diameters. The analysis indicated a main effect of round type ($F(1, 31) = 22.97$, $p < .001$), a main effect of block ($F(9, 279) = 23.91$, $p < .001$) as well as an interaction between the two factors ($F(9, 279) = 15.42$, $p < .001$). Further analyses indicated that pupil diameters on TFT and non-TFT rounds were not different during the first two blocks (p's $> .6$). During the third block, pupil diameters were greater on non-TFT rounds than on TFT rounds ($t(31) = 2.04$, $p = .05$), and this difference remained until the end of the outcome display (p's $< .001$).

To more precisely examine the onset of the differential pupil responses on TFT and non-TFT rounds, we next analyzed pupil diameters at every time point throughout the outcome display (Figure 9). Specifically, we compared pupil diameters at every time point on TFT rounds and non-TFT rounds. Paired t-tests indicated that pupil diameters were greater on non-TFT rounds than on TFT rounds from 520 milliseconds onward (p's $< .05$).

As for cooperation, participants cooperated on 41% of the rounds (SD = .22). We divided the sequence into five blocks of 20 rounds. As with the experiments reported above, our analysis indicated that cooperation rates decreased over time ($F(4, 124) = 7.47$, $p < .001$).

Discussion

Experiment 4 found that pupil diameters were greater when an otherwise reciprocative opponent failed to reciprocate than when the opponent reciprocated. Furthermore, such pupil responses emerged approximately 520 milliseconds after participants were presented with the outcome of a round, and this response remained for at least another 1.5 seconds without obvious decay. Such a finding provides converging evidence that is consistent with our results using the tone probe paradigm. Participants appear to develop expectations about reciprocation during repeated interactions with a reciprocating opponent and engaged in greater cognitive processing when such expectations were violated. Moreover, our findings using pupillometry provide additional information to the temporal course of such cognitive processing and alleviate any concern with the overt nature of the tone probe paradigm.

General Discussion

Four experiments examined participants' expectations about reciprocation during an iterated PD. In all the experiments, participants played with a TFT opponent who occasionally failed to reciprocate and instead did the opposite of what a reciprocating player would do. Experiment 1 found that participants were slower to respond to tones played on non-TFT rounds than on TFT rounds, suggesting that they engaged in greater cognitive processing when the otherwise reciprocative opponent did not reciprocate. This finding is consistent with the idea that participants developed expectations about reciprocation and such expectations were violated on non-TFT rounds. Experiment 2 again found that participants exhibited differential responses to tones during on TFT and non-TFT rounds, but only when reciprocation was associated with shorter delays. When reciprocation was associated with longer delays no such sensitivity was observed. Such finding suggests that delaying reciprocation hurts cooperation (Liu & Luhmann, 2014a) by attenuating participants' expectations about their partners' objectively predictable behavior. Experiment 3 further examined whether participants' expectations focus more on opponents' behavior or on their own payoffs. Our results suggest that participants' expectations concern both aspects of the interaction. Experiment 4 utilized pupillometry to track the cognitive processing as the outcomes of interactions were revealed. Our results indicated non-TFT rounds elicited increased pupil diameter relative to TFT rounds. This finding, which is consistent with the tone probe findings, suggests that participants engaged in greater processing when their opponents did not reciprocate. Furthermore, such pupil responses began approximately 500 milliseconds after outcome information was presented.

Three experiments in the current study utilized a concurrent tone-discrimination paradigm to monitor cognitive processing during iterated PDs. This paradigm was previously used to explore cognitive processing during contingency learning (Liu & Luhmann, 2013, in press). Our current findings in PDs are consistent with previous findings that unexpected information or events compromises performance in the secondary tone-discrimination task. Both the previous and current findings are consistent with the idea that processing of unexpected information recruits more cognitive resources than processing of expected information.

As mentioned in the introduction, the TFT strategy has been demonstrated to encourage cooperation (e.g., Bakers & Rachlin, 2001; Silverstein et al., 1998; Fantino et al., 2006; Kuhlman & Marshello, 1975; Stephens et al., 2002, 2006; Yi et al., 2005) but relatively little is known about why individuals learn to cooperate with a reciprocative opponent like TFT. Findings from the current study suggest that the way people learn to cooperate in iterated PDs shares many characteristics with traditional associative learning. Specifically, many models of associative learning suggest that learning relies on the computation of prediction error (e.g., Catena et al., 1998; Danks et al., 2003; Hogarth & Einhorn, 1992; Luhmann & Ahn, 2007; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972), which is the difference between learners' predictions about future events and what actually happens. The patterns we observed in responses to tones and pupil diameters are consistent with this idea. Specifically, when playing against a TFT opponent repeatedly, participants may have developed expectations about reciprocation. During the minority of rounds on which the TFT opponent did not reciprocate, these expectations about reciprocation would have been violated, which should have led to greater prediction errors (unlike TFT rounds, on which the TFT opponent behaved as expected). In those unexpected situations, greater cognitive processing would have been evoked, consistent with the fact that participants showed slower responses to concurrent tone probes as well as increased pupil size.

One objective of the current study was to characterize how individuals learn about reciprocation: whether they do so by associating their own previous behavior with their opponent's current behavior or merely by associating their own previous behavior with their own current payoffs. Our results suggest that participants learn both associations equally well. Indeed, these two associations are highly redundant in standard PD games, and our findings could suggest there is simply no difference between the two associations. However, there is reason to distinguish between these two associations. Specifically, these two learning pathways mirror two alternative learning strategies described in the literature of reinforcement learning: model-free vs. model-based reinforcement learning (Sutton & Barto, 1998). In model-free learning, individuals simply associate actions with their ultimate value or reward and choose actions that are associated with higher rewards. According to this approach, successfully learning about reciprocation in an iterated PD involves cooperation becoming more valuable because it is associated with higher future payoffs. In model-based learning, individuals acquire an internal model that describes how taking one action may change the nature of future choices, either in which alternatives are available or in what rewards those actions will produce. Under model-based learning, agents take actions by "simulating" the consequences of the current choice and all future choices that follow. According to this approach, successfully learning about reciprocation in an iterated PD involves understanding that one's cooperation leads to the future cooperation of the reciprocating opponent, and that such cooperation is what ultimately predicts higher payoffs. There has been evidence that these two kinds of learning rely on different neural substrates (e.g., Doll, Shohamy, & Daw, 2014; Glascher, Daw, Dayan, & O'Doherty, 2010), require cognitive control to a different extent (Otto, Gershman, Markman, & Daw, 2013), and are not equally subjective to extinction (e.g., Tricomi, Balleine, & O'Doherty, 2009). One possibility is that the lack of differences reported in Experiment 3 may be a result of the methodology used. Employing more sensitive measurements, such as was evident with pupillometry in Experiment 4, may provide better chance to observe potential differences between model-free and model-based mechanisms of learning in an iterated PD.

Reference

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Baker, F., & Rachlin, H. (2001). Probability of reciprocation in repeated prisoner's dilemma games. *Journal of Behavioral Decision Making, 14*, 51–67.

Baker, F., & Rachlin, H. (2002). Self - control by pigeons in the prisoner's dilemma. *Psychonomic Bulletin & Review, 9*, 482–488.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–162). Cambridge, MA: Cambridge University Press.

Duffy, S., & Smith, J. (2012). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? MPRA Paper No. 35906, LMU.

Fantino, E., Gaitan, S. C., Meyer, S., & Stolarz-Fantino, S. (2006). Self-control in a prisoner's dilemma game: The role of trial spacing and an interfering task. *Journal of Behavioral Decision Making, 19*, 273-284.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*, 289-344.

Komorita, S. S., Hilty, J. A., & Parks, C. D. (1991). Reciprocity and cooperation in social dilemmas. *Journal of Conflict Resolution, 35*, 494–518.

Kuhlman, D. M., & Marshello, A. F. J. (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology, 32*, 922-931.

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science, 7*, 18-27.

Liu, P., & Luhmann, C. C. (2013). Evidence that a transient, but cognitively demanding process underlies forward blocking. *Quarterly Journal of Experimental Psychology, 66*, 744-766.

Liu, P., & Luhmann, C. C. (2014a). Cognitive Constraints on Cooperation in an Iterated Prisoner's Dilemma. Unpublished Manuscript.

Liu, P., & Luhmann, C. C. (2014b). Evidence for On-Line Processing during Causal Learning. *Learning and Behavior.*

Locey, M. L., & Rachlin, H. (2012). Temporal dynamics of cooperation. Journal of *Behavioral Decision Making, 25*, 257-263.

Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences, 95*, 13755–13758.

Rapoport,A., & Chammah,A.M. (1965). *Prisoner's dilemma: A study in conflict and cooperation.* Ann Arbor: University of Michigan Press.

Renner, K. E. (1964). Delay of reinforcement: A historical review. *Psychological Bulletin, 61*, 341-361.

Parks, C. D., & Rumble, A. C. (2001). Elements of reciprocity and social value orientation. *Personality and Social Psychology, 27*, 1301–1309.

Rachlin, H., Brown, J., & Baker, F. (2000). Reinforcement and punishment in the prisoner's dilemma game. In: D. Medin (Ed.), *The psychology of learning and motivation, 40* (pp. 327–364). New York, NY: Academic Press.

Roth, A. E. (1995). Introduction. In J. H. Kagel, & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 3–20). Princeton, NJ: Princeton University Press.

Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., & Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *Neuroimage, 37*, 1017–1031.

Silverstein, A., Cross, D., Brown, J., & Rachlin, H. (1998). Prior experience and patterning in a prisoner's dilemma game. *Journal of Behavioral Decision Making, 11*, 123-138.

Stevens, J. R., Cushman, F. A., & Hauser, M. D. (2005). Evolving the psychological mechanisms for cooperation. *Annual Review of Ecology, Evolution, and Systematics, 36*, 499–518.

Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences, 8*, 60–65.

Stephens, W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an iterated prisoner's dilemma. *Science, 298*, 2216–2218.

Stephens, W., McLinn, C. M., & Stevens, J. R. (2006). Effects of temporal clumping and payoff accumulation on impulsiveness and cooperation. *Behavioral Processes, 71*, 29-40.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: The MIT Press.

Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (In press). Better late than never? The effect of feedback delay on ERP indices of reward processing, *Cognitive, Affective, and Behavioral Neuroscience.*

Yi, R., Johnson, M. W., & Bickel, W. K. (2005). Relationship between cooperation in an iterated prisoners dilemma game and the discounting of hypothetical outcomes. *Learning & Behavior, 33*, 324–336.

Figure 1. The payoff matrix detailing the Prisoner's Dilemma used in all proposed experiments. For example, when a player defects and her opponent cooperates, the player receives six points and the opponent receives one point, as indicated in the top, right cell.

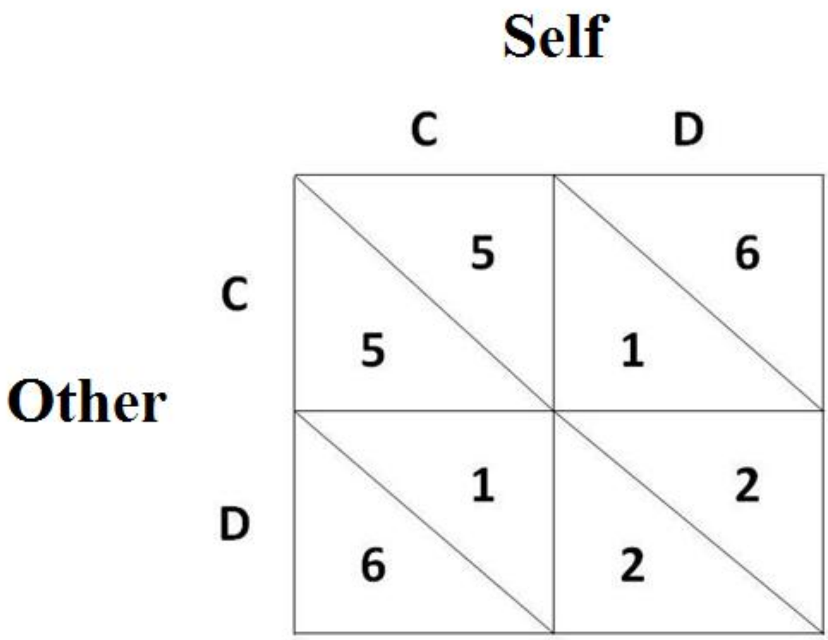

Figure 2. The procedure of each round of the iterated PD used Experiments 1, 2, and 4.

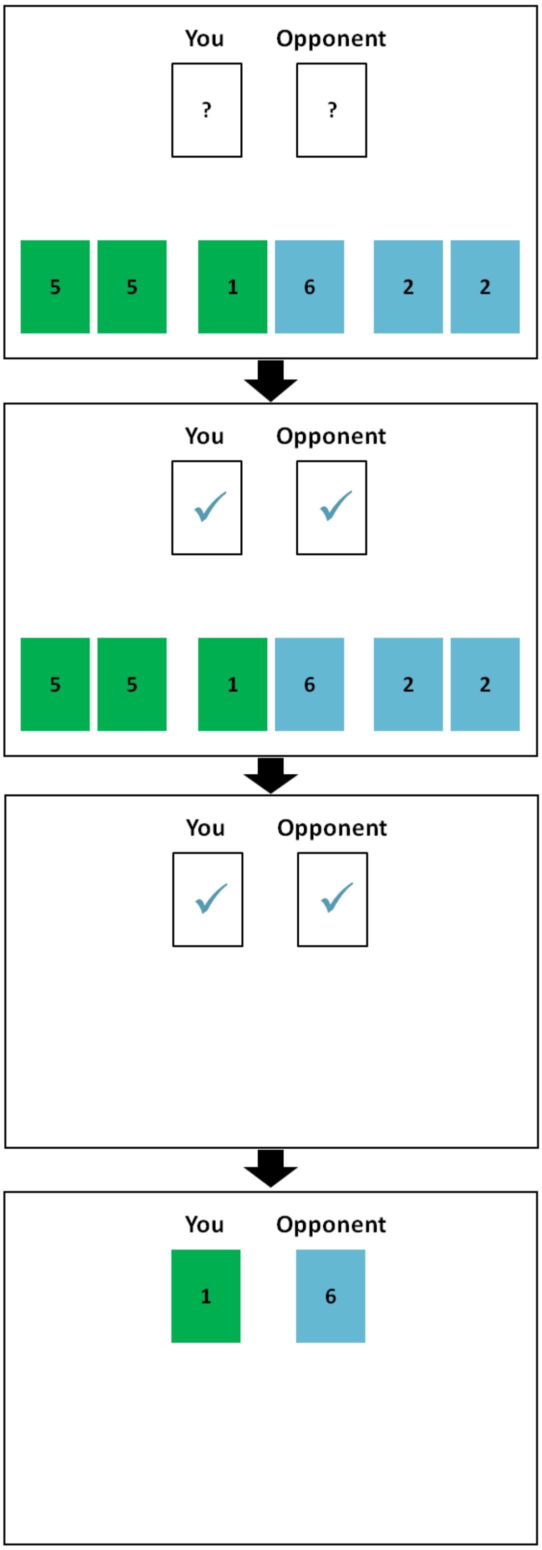

Figure 3. Response time (RT) to the tones played on TFT rounds and on non-TFT rounds in Experiment 1. As expected, responses on non-TFT rounds were slower than responses on TFT rounds.

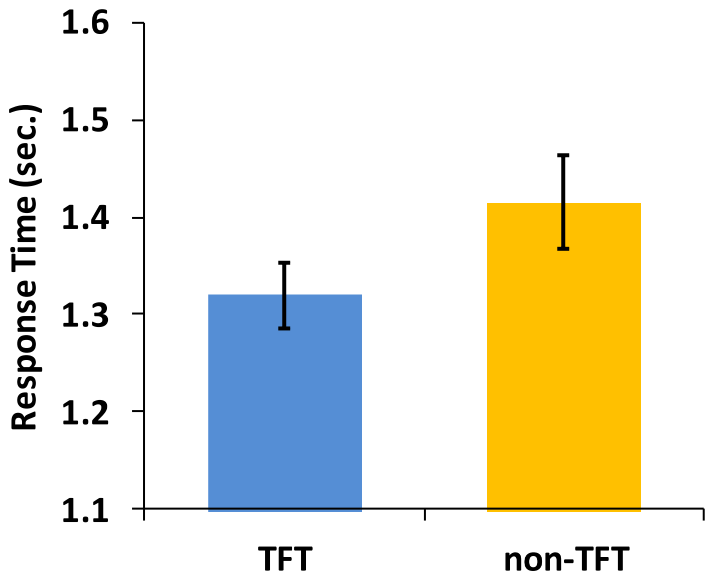

Figure 4. RTs to the tones played on TFT rounds and on non-TFT rounds in the Long-ITI and Short-ITI conditions Experiment 2. In the Short-ITI condition, responses on non-TFT rounds were slower than responses on TFT rounds. No such difference was observed in the Long-ITI condition.

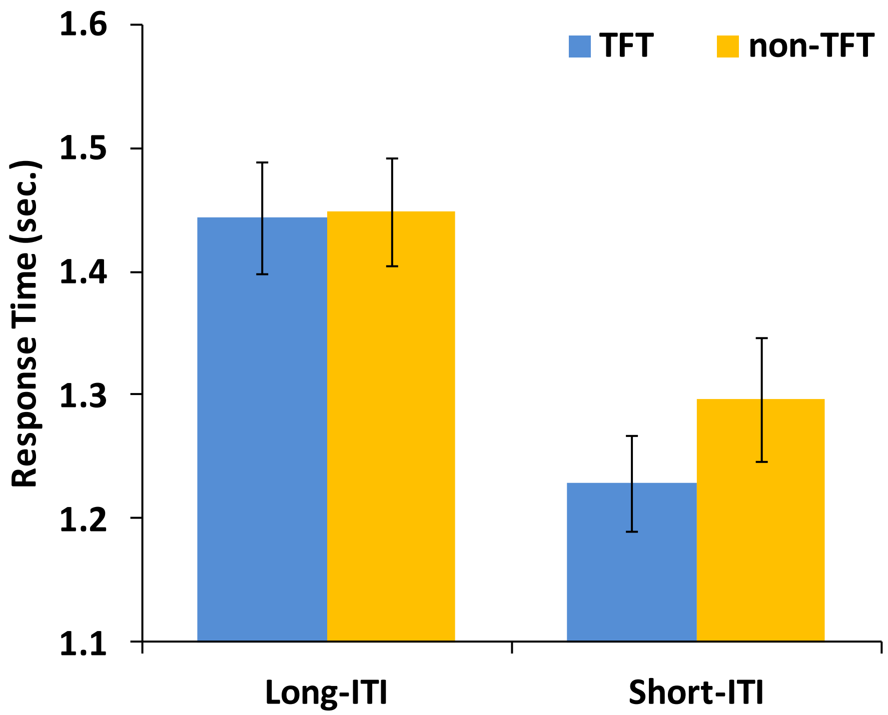

Figure 5. Cooperation rates over time in the Long-ITI and Short-ITI conditions in Experiment 2. Cooperation decreased over time in both conditions and was not different between the two conditions.

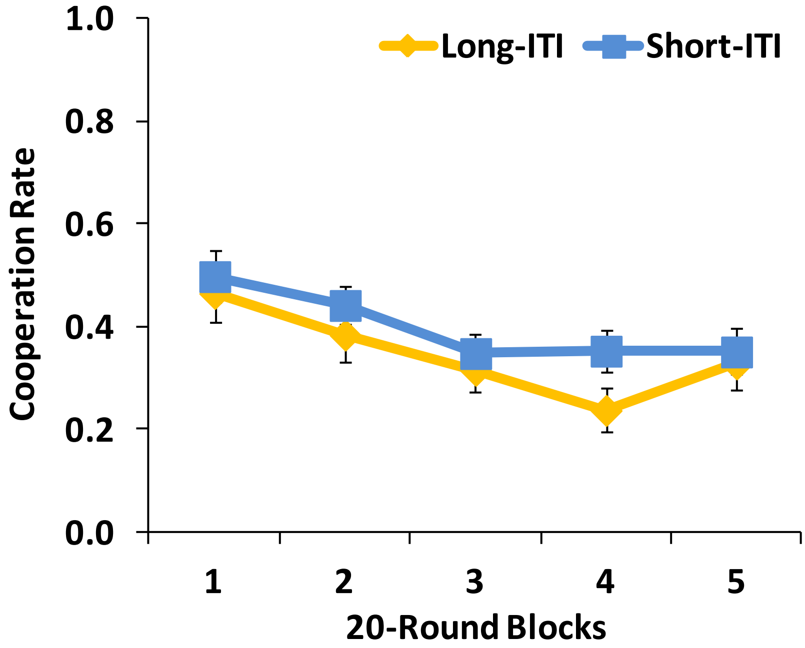

Figure 6. Illustration of the outcome display on the focus rounds in Experiment 3. On the left is an example from the Choice-Focus condition and on the right is an example from the Payoff-Focus condition. The round preceding the outcome display is identical to that illustrated in Figure 2.

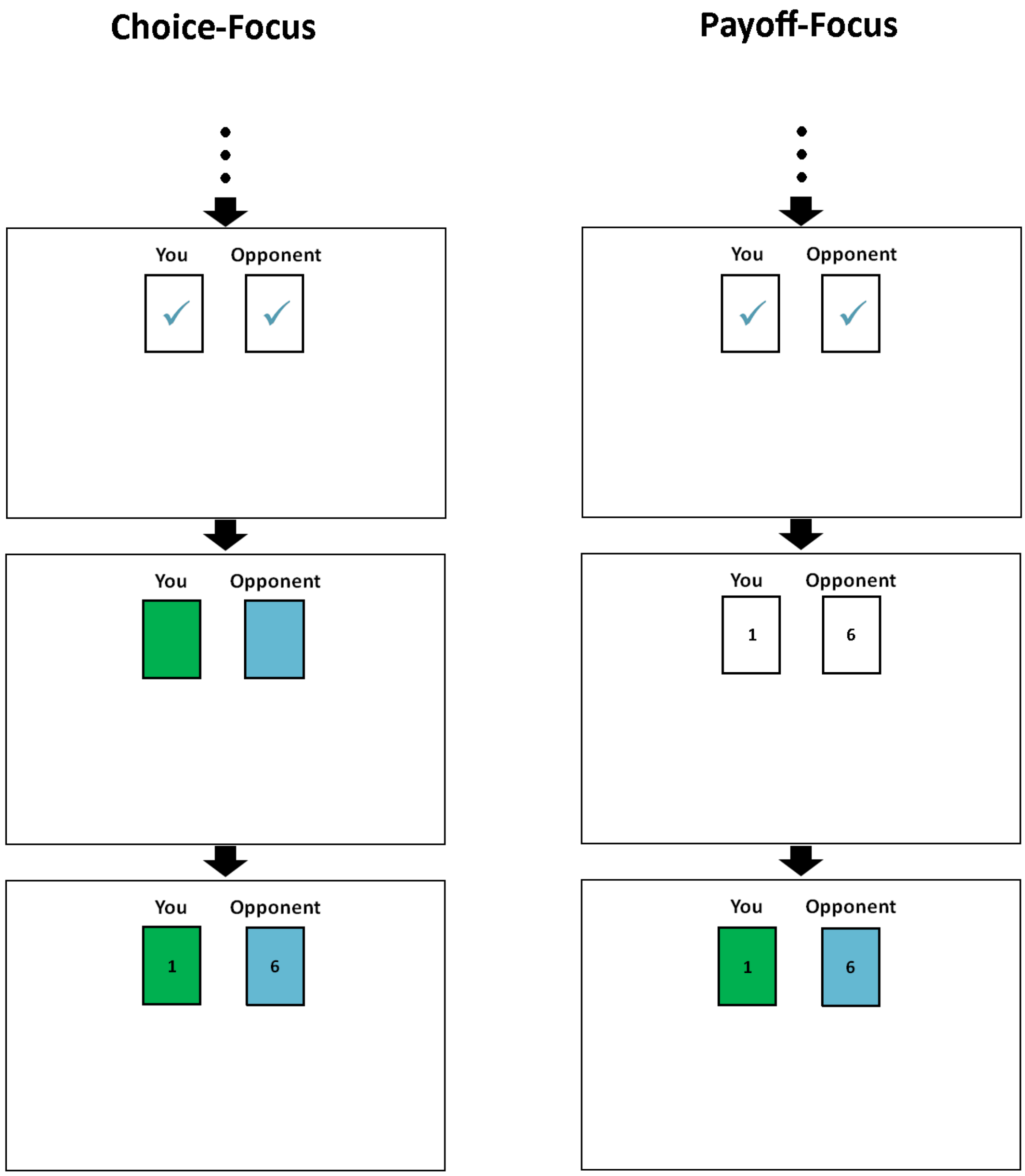

Figure 7. RTs to tones played on TFT and non-TFT rounds in the Choice-Focus and Payoff-Focus conditions in Experiment 3. Responses on non-TFT rounds were slower than responses on TFT rounds in both conditions. The magnitudes of the RT effects were not different in the two conditions.

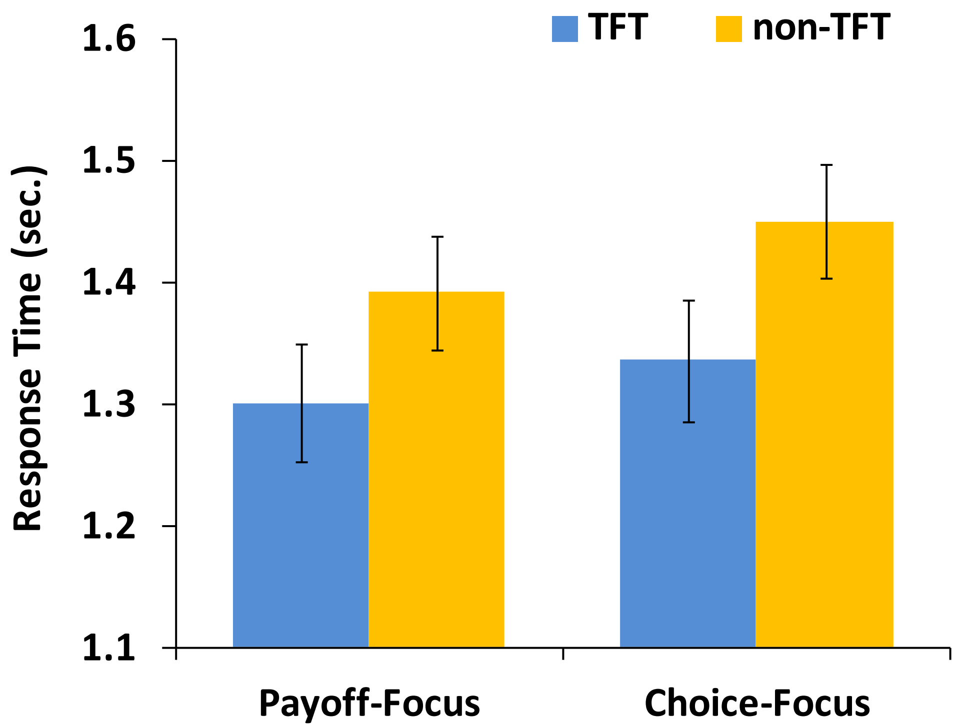

Figure 8. Cooperation rates over time in the Choice-Focus and Payoff-Focus conditions in Experiment 3. Cooperation decreased over time in both conditions and was not different between the two conditions.

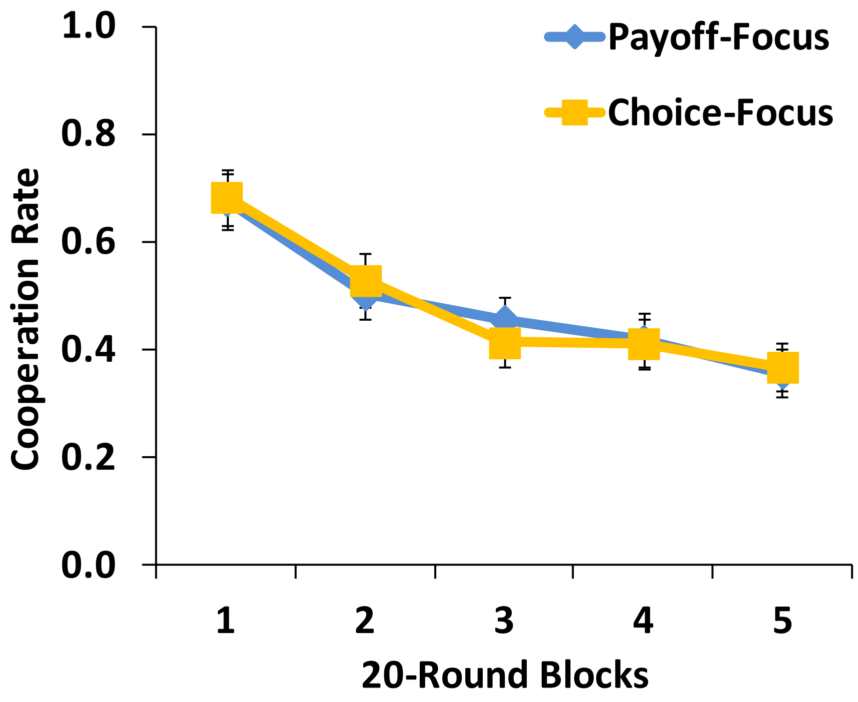

Figure 9. Pupil diameters during baseline and during outcome display on TFT and non-TFT rounds in Experiment 4. “0” on X-axis represents the beginning of outcome display. From 520 milliseconds onwards, pupil diameter on non-TFT rounds was greater than pupil diameter on TFT rounds.