Indirect Higher Order Beliefs and Cooperation

WU, JIABIN

Department of Economics, University of Oregon

18 February 2016

Indirect Higher Order Beliefs and Cooperation

Jiabin Wu*

Department of Economics, University of Oregon

515 PLC, 1285 University of Oregon, Eugene, OR 97403

February 18, 2016

Abstract

This paper experimentally examines why communication may matter for inducing cooperation in strategic interactions involving intermediaries. We consider a three-player centipede game in which the first and the third players do not interact sequentially, but only through the second player. We posit that the third player's decision to cooperate depends on his *indirect higher order belief*, that is, his belief about what the first player believes the second player would choose. The evidence demonstrates that communication between the first and the third player can effectively induce cooperation from the third player through shaping his indirect higher order belief.

Keywords: indirect higher order beliefs, communication, psychological game theory, guilt aversion, sequential reciprocity, social preferences, behavioral economics, experimental economics

JEL Code: C72, C91, D03, D83.

*Email: jwu5@uoregon.edu, Phone: (541) 346-5778. The author is indebted to Noah Lim for his continuous support, guidance and encouragement on this project. The author is grateful to Pierpaolo Battigalli, Gary Charness, Martin Dufwenberg, Joshua Miller, Aldo Rustichini, William Sandholm, Itai Sher and participants from 2015 ESA North American Meeting for their helpful comments. The author thanks Kristian Chavira and Michael Matuszak for excellent research assistance.

1 Introduction

Pioneered by Geanakoplos, Pearce and Stacchetti (1989), psychological game theory has become increasingly important in explaining behavior in strategic interactions. The theory argues that the motivations of decision makers, who may be influenced by psychological drivers such as guilt aversion or reciprocity, depend on their beliefs about others' behavior as well as the beliefs of others. This theory has been shown to provide a powerful framework for understanding various real-world phenomena in which individuals deviate from standard economic predictions.[1] For example, Charness and Dufwenberg (2006) experimentally show that a particular form of communication, promises, effectively foster bilateral partnerships through influencing the players' higher order beliefs.

From a theoretical point of view, the belief structures that determine the players' motivations in psychological game theory become richer and more complex as the number of players in the game increases (see Battigalli and Dufwenberg (2007, 2009)): In a two-player game, a player's action depends on what he believes about the other player's belief about his own action. We call this type of beliefs the *direct* higher order beliefs. When there are three players, in addition to direct higher order beliefs, a player's action may also depend on his belief about another player's belief about the third player's action. We call this type of beliefs the *indirect* higher order beliefs.

Indirect higher order beliefs may be important in strategic settings in which some parties do not interact sequentially but through intermediaries.[2] For example, in a company, a worker may not interact directly with the CEO but only through supervisors. How a worker responds to the CEO's decision to appoint a supervisor depends on what he believes about the CEO's purpose for doing so (for example, the supervisor may either choose to help the worker to be more productive, or monitor him more closely which increases his disutility). If the worker perceives the supervisor's action to be fully controlled by the CEO, then the worker can interpret the supervisor's action to be the CEO's intention. However, if the supervisor has some autonomy, then the worker's belief

[1] Important theoretical contributions to psychological game theory include Koplin (1992), Rabin (1993), Dufwenberg and Kirchsteiger (2004), Battigalli and Dufwenberg (2007, 2009), Dufwenberg et al (2011) and Battigalli et al (2014). See also general surveys by Dufwenberg (2008) and Attanasi and Nagel (2008). Experimental tests of this theory include Güth et al (1994), Berg et al (1995), Glaeser et al (2000), Dufwenberg and Gneezy (2000), Hannan et al (2002), Charness and Rabin (2002, 2005), Charness and Dufwenberg (2006, 2011) and Bartling and Fischbacher (2012).

[2] Intermediation is commonly observed in inter-firm interactions such as outsourcing (Hamman et al (2010), Coffman (2011)), in political activities such as bribery (Drugov et al (2014)), in workplace delegation tasks (Bartling and Fischbacher (2012)).

about the CEO's belief about the supervisor's action (an indirect higher order belief) may play a greater role in affecting the worker's response. In this case, direct communication between the CEO and the worker may effectively shape the worker's indirect higher order belief and help enhance productivity in the workplace.

So far, the experimental literature on psychological game theory has only focused on examining the impact of communication on behavior through influencing direct higher order beliefs in two player games. To the best of our knowledge, no previous work has examined how communication may affect behavior through influencing indirect higher order beliefs in games with more than two players. This paper serves as an attempt to address this issue.

We first present a simple model based on psychological game theory to flesh out how indirect higher order beliefs can influence players' behavior. We introduce a three-player centipede game in which the first player interacts with the third player only through the second player. When the first player considers his moves, he knows that the second player can either choose to hurt the third player while benefiting both the first and second players, or to allow the third player to make a decision. If the third player gets to make a decision, he can either choose a cooperation outcome which generates high payoffs for the first two players; or to hurt them while benefiting himself. The decision made by the third player may be influenced by his higher order beliefs. If the third player chooses to benefit himself at the expense of the first player, he may experience disutility if he believes that the first player had intended to achieve the cooperation outcome. In our model, the third player's belief about the first player's expectation about the cooperation outcome is not only captured by his belief about the first player's belief about his probability of choosing the cooperation outcome (i.e., his direct higher order belief), but more critically, by his belief about the first player's belief about the second player's action (i.e., his indirect higher order belief). Specifically, if the third player believes that the first player expects with high probability that the second player is going to choose to hurt the third player, then the third player must believe that the first player's expectation of attaining the cooperation outcome is low. In this case, the third player would experience less (or even no) disutility if he chooses to benefit himself at the expense of the first player. Naturally, our model also allows the third player to suffer disutility from hurting the second player if he believes that the second player believes that he will choose the cooperation outcome.

Based on the insights from the model, we experimentally examine if communication can foster cooperation in this three-player centipede game through indirect higher order beliefs. We begin by allowing the first player to send a free-form message to the third player before the game starts.

We find that this indeed changes the indirect higher order belief of the third player and also the frequency of cooperation outcomes, compared to the baseline treatment without communication. We further use a treatment where the message that the first player could send to the third player was predetermined and contained information about the first player's belief about the second player's behavior. The rationale for this treatment is as follows: given that communication can affect beliefs, a message that explicitly indicates the first player's belief about the second player's behavior would shift the third player's indirect higher order belief and the frequency of cooperation outcomes even more. Our experimental results confirm this prediction.

This paper is also closely related to the experimental literature on delegation.[3] The literature suggests that the act of delegation may alter some parties' perceptions about others in strategic interactions. In particular, Bartling and Fischbacher (2012) study a game that shares some similarities with ours, in which a dictator decides between a fair and an unfair allocation, or to delegate the decision right to a delegee. The dictator's and the delegee's interests are aligned and they both benefit from the unfair allocation. Two receivers are hurt by the unfair allocation and they can punish the dictator and/or the delegee. The paper finds that punishment can be effectively shifted to the delegee if the dictator delegates the decision right to the delegee and the delegee makes the unfair choice. The authors also elicit indirect higher order beliefs (i.e., the receivers' beliefs about the dictator's belief about the delegee's choice) from the receivers and they argue that it is possible that the action of delegation can effectively influence the receivers' punishment decision through influencing their indirect higher order beliefs. Our paper differs from Bartling and Fischbacher (2012) in two ways: First, in the centipede game we study, the first player cannot shift the blame to the second player through delegation and consequently avoid punishment. Hence, our game captures situations where delegation is unable to influence the third player's indirect higher belief. Second, we show that communication between the first and the third player serves as an effective way to foster cooperation through influencing the third player's indirect higher order belief in such situations.

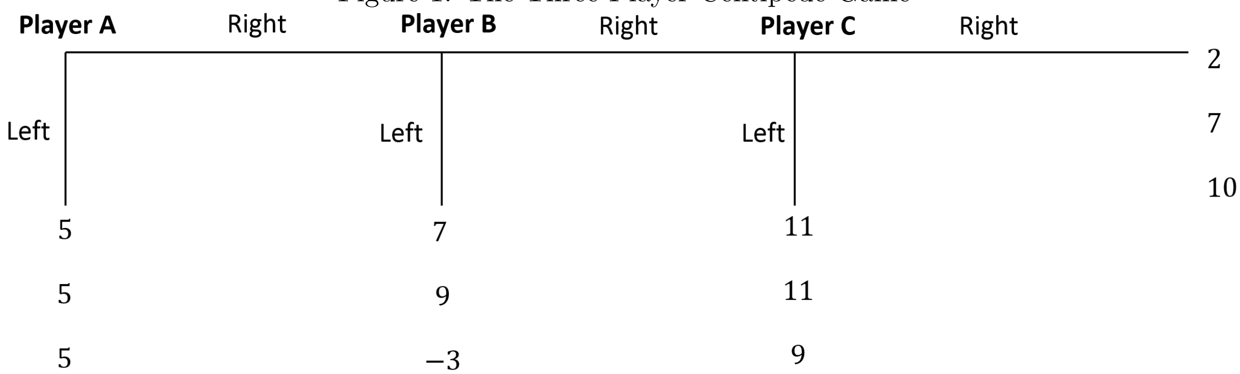
The paper is organized as follows. Section 2 introduces the three-player centipede game we study. Section 3 presents a simple model based on psychological game theory that explains the role of indirect higher order beliefs in influencing behavior. Section 4 experimentally investigates how different forms of communication could affect the players' beliefs and behavior. Section 5 concludes.

[3]See Hamman et al (2010), Coffman (2011), Bartling and Fischbacher (2012), Charness et al (2012), Lai and Lim (2012) Erat (2013), Fehr et al (2013), Oexl and Grossman (2013), Bartling, Engl and Weber (2014), Bartling, Fehr and Herz (2014) and Drugov et al (2014), among others.

2 The Three-Player Centipede Game

Figure 1 presents the three-player centipede game we design to examine how indirect higher order beliefs may affect behavior. This game captures situations in which two people do not interact sequentially, but through an intermediary.

Figure 1: The Three-Player Centipede Game



In the game shown in Figure 1, all the three players have two actions: Left or Right. If Player A chooses Left, all the three players receive payoffs of 5. If Player A chooses Right, then Player B gets to choose between Left or Right. If Player B chooses Left, Player A gets 7, Player B gets 9 and Player C receives a negative payoff of -3. On the other hand, if Player B chooses Right, then Player C can choose between Left or Right. If Player C chooses Left, both Players A and B get 11 each and Player C gets 9. If Player C chooses Right instead, he gets a higher payoff of 10, but Player A gets only 2 and Player B gets only 7.

Under the self-interested assumption and applying backward induction, it is straightforward to see that the unique subgame perfect equilibrium (SPE) is (Right, Left, Right) and the corresponding equilibrium outcome is (7, 9, -3). Note that (Right, Right, Left) corresponds to the outcome (11, 11, 9), which payoff dominates the SPE outcome. We call this outcome the Cooperation outcome.

How might the players achieve the Cooperation outcome (supposing they wish to)? To answer this, it is useful to examine the key features of the game: First, Player A cannot directly "coordinate" with Player C on the Cooperation outcome by choosing Right. Whether he can do so is also determined by Player B's actions. Second, Player B can benefit himself while hurting Player C significantly. Moreover, if Player A chooses Right, Player A also benefits from Player B's action that hurts Player C. Hence, whether Player A's intent was to reach the Cooperation outcome or to benefit himself while hurting Player C through Player B is ambiguous to Player C. This remains the case even when Player C observes Player B choosing Right. Third, If Player B chooses Right, Player C can either achieve the Cooperation outcome or hurt both Players A and B, with Player A being hurt more severely.

From this game, one can readily see that the key to achieving the Cooperation outcome hinges on Player C's belief about Player A's intention. Then what might capture Player C's belief about Player A's intention? Applying psychological game theory, we argue that Player C's indirect higher order belief, specifically, his belief about Player A's belief about Player B's action, plays a key role. If Player C believes that Player A believes that Player B has a high probability of choosing Left, then he believes that Player A's intention in choosing Right is to reap higher payoffs at his expense. Conversely, if Player C believes that Player A thinks that Player B has a high probability of choosing Right, then he believes that Player A's action of choosing Right is to reach the Cooperation outcome. We now proceed to formalize this intuition.

3 A Simple Model Based on Psychological Game Theory

We present a simple model based on psychological game theory to illustrate how Player C's indirect higher order belief may affect his decision in the three-player centipede game above. To keep the model simple, we assume that the disutility that Player C experiences when he hurts Players A and B can be captured by guilt aversion (Battigalli and Dufwenberg (2007)).[4] It is important to note that we do not claim that guilt aversion is the only psychological driving force behind Player C's action. The main focus of this paper is to examine how the indirect higher order beliefs affect behavior. In Appendix A, we provide an alternative model that assumes that Player C exhibits reciprocity (Dufwenberg and Kirchsteiger (2004)) and show that our main insights do not change.

We begin by establishing a system of notations for describing beliefs in our model: Let $\delta_{iC} \in [0, 1]$ denote player i's belief about Player C choosing Left, $i \in \{A, B\}$. This is the first order belief that a player assigns to Player C's decision. A player can also form beliefs on the first order belief (with certain probability measure). Let $\delta_{ijC} \in [0, 1]$ denote the mean of player j's belief about player i's belief about Player C choosing Left, $i \in A, B$, $j \in \{A, B, C\}$. This is the higher order belief that a player assigns to Player C's decision. Similarly, let $\tau_{iB} \in [0, 1]$ denote player i's belief about

[4] Guilt aversion is a well-established psychological driver of behavior. Studies in both social psychology (see Baumeister, Stillwell and Heatherton (1994, 1995)) and neuroscience (Chang, Smith, Dufwenberg and Sanfey (2011)) provide evidence of guilt aversion. Applications include Charness and Dufwenberg (2006, 2011), and Chen and Lim (2013) who show that guilt aversion drives effort choices in team contests.

Player B choosing Left, $i \in \{A, C\}$. This is the first order belief that a player assigns to Player B's decision. Let $\tau_{jiB} \in [0, 1]$ denote the mean of player j's belief about player i's belief about Player B choosing Left, $i \in \{A, C\}, j \in \{A, B, C\}$. This is the higher order belief that a player assigns to Player B's decision.

Next, we differentiate between two types of higher order beliefs. The first type is the *direct* higher order belief, which captures a player's belief about another player's belief about his *own* action. In the context of our three-player game, the beliefs τ_{BAB}, δ_{CBC} and δ_{CAC} are examples of direct higher order beliefs. The other type is the *indirect* higher order belief, which captures a player's belief about another player's belief about a third player's action. In this paper, our primary interest is on Player C's belief about Player A's belief about Player B's actions. The formal notation for this indirect higher order belief is τ_{CAB}.

Consider Player C's decision after both Players A and B chose Right. Choosing Right yields Player C a higher material payoff ($10 > 9$), but he may experience guilt because this action reduces the payoffs of Players A and B relative to what they expect to receive. Let θ denote Player C's degree of guilt aversion towards the other two players.

We begin by modeling the guilt aversion that Player C may feel towards Player B. Our logic is similar to that in the two-player guilt aversion model in Charness and Dufwenberg (2006): Since Player C gets to make a choice only if Player B chooses Right, his guilt aversion towards B would simply be driven by how much he hurts Player B relative to how much Player B expects to get. In our game, $4\delta_{CBC} = 11\delta_{CBC} + 7(1 - \delta_{CBC}) - 7$ is the payoff loss that Player C believes that Player B believes that she herself (Player B) would incur if he (Player C) chooses Right. Therefore, the term $\theta \times 4\delta_{CBC}$ captures the psychological loss that Player C experiences towards Player B if he chooses Right.

Next, we describe the guilt aversion that Player C may feel towards Player A. We assume that Player C feels more guilty towards Player A the higher $(1 - \tau_{CAB})$ becomes. This is because when the indirect higher order belief τ_{CAB} is high, Player C believes that Player A chose Left because Player A intended to obtain the subgame perfect equilibrium outcome, in which Player C is hurt badly. On the contrary, when τ_{CAB} is low, Player C believes that Player A expects him to make a decision. Moreover, it is also reasonable to assume that he thinks that the reason why Player A chooses Right is that Player A believes that he is likely to choose Left (i.e. high δ_{CAC}). Hence, the direct higher order belief δ_{CAC} may be inversely related to τ_{CAB} when τ_{CAB} is low.

The payoff loss Player C believes that Player A would incur when he chooses Right is: $11\delta_{CAC} + 2(1 - \delta_{CAC}) - 2 = 9\delta_{CAC}$. Therefore, $\theta \times (1 - \tau_{CAB}) \times 9\delta_{CAC}$ captures the psychological loss of

Player C from hurting Player A when he chooses Right.

In summary, the utilities of Player C from choosing Left and Right are given by

$$U_C(Left) = 9; \tag{1}$$

$$U_C(Right) = 10 - \theta(4\delta_{CBC} - (1 - \tau_{CAB})9\delta_{CAC}). \tag{2}$$

Given this, Player C chooses Left if

$$9 > 10 - \theta(4\delta_{CBC} - (1 - \tau_{CAB})9\delta_{CAC}). \tag{3}$$

When τ_{CAB} is low (so that δ_{CAC} should be high as well) and δ_{CBC} is high, i.e., Player C believes that both Players A and B expect to reach the Cooperation outcome, Player C would also choose Left.

Before we proceed to the lab experiments, we present two remarks regarding the model: First, similar to Charness and Dufwenberg (2006), we do not model Player A's and B's decisions explicitly and we do not employ equilibrium concept.[5] This is because our primary focus is on how Player C's indirect higher order belief affects his decision. More importantly, given the unique structure of the centipede game, Player C can only make a choice after both A and B have chosen Right. This means that from Player C's point of view, he does not observe variation in A and B's choices. Hence, any guilt he feels towards the other two players stems from his beliefs about the payoff discrepancies between what the other two players had planned to achieve, and what they would actually receive. These beliefs are independent of the modeling choice of Player A's and B's preferences, as well as whether the players have common knowledge of others' preferences.

Second, we allow Player A to send a message to Player C in the experiments. However, we do not think it is necessary to explicitly model communication as a strategic choice of Player A. Player C moves at the last stage of the centipede game. If communication indeed operates as a strategic choice of Player A in the game, it is because of Player C's belief-dependent utility, such that communication can change Player C's behavior through the changes it induces in his beliefs.[6]

[5]Note that Charness and Duwenberg (2006) simply assumes that the first mover in the two-player trust game is self-interested.

[6]See Battigalli et al (2014) for a discussion on incorporating cheap talk into psychological game theory for a 2-person game.

4 Experimental Tests

The model above illustrates the channel through which the indirect higher order belief τ_{CAB} may affect Player C's decision. Based on the insights of the model, we conduct a set of laboratory experiments to investigate if communication induces cooperation in the three-player centipede game through influencing τ_{CAB}.

The experiments were conducted with undergraduate students at the University of Wisconsin-Madison. There were 4 treatments, with each treatment consisting of either two or three sessions. In each session, there were 18 subjects and 12 decision rounds (with 3 practice rounds before the first round started). Each session lasted about one hour. In each round, subjects were randomly and anonymously matched into 6 groups of 3 players each. After each round, subjects were re-matched in the same manner by the computer program used to implement the experimental session. We rotated the roles so that each subject had the opportunity to be Player A, B and C.[7] Each subject participated in only one session, and the average earnings were around $15.

4.1 Treatments

The main objective of the four treatments is to examine if communication changes Player C's indirect higher order belief, consequently affecting behavior and the game outcomes. The details of the four treatments are as follows.

1. The first treatment T1 (2 sessions) was the baseline treatment where players did not communicate with one another.

2. The second treatment T2 (2 sessions) is called the *AtoC Free-Form Private Treatment*. In this treatment, Player A could send any message he wishes (or none at all) to Player C prior to choosing Left or Right. All players knew that Player A could send a message to Player C. However, information about whether Player A chose to send a message and the content of the message were not revealed to Player B. In this way, the message was private only to Players A and C.

3. The third treatment T3 (3 sessions) is called the *AtoC Free-Form Public Treatment*. This treatment is similar to T2 with the following exception: After Player A's message was sent, it

[7] Specifically, we rotated the roles so that each participant plays each of the three roles exactly 4 times. However, we did not inform the participants that they would do so in the experimental instructions. In the statistical tests, we adjust for potential within-subject correlation by clustering at the individual subject level.

would be shown on the screens of both Players B and C. In this way, the message was public to all three players.

4. The fourth treatment (T4) (2 sessions) is called the *AtoC Fixed Message Treatment*, in which the message that Player A could send to Player C was predetermined. In this treatment, Player A can choose between two messages to send to Player C prior to choosing Left or Right.

   - Message 1: ***I intend to choose <u>RIGHT</u>, because I believe that Player B will choose <u>RIGHT</u> as well.***
   - Message 2: ***Blank message.***

   Once a message was sent, it would be shown to Player C. In this treatment, Player B could not tell if Player A had sent the message or not, but he knew that Player A had these communication options.

Note that the nature of communication in the experiment is "forward directed" in the sense that the message is sent by one player to another player who moves after him. This feature contrasts with experimental studies that employ "backward directed" communication (e.g. Charness and Dufwenberg (2006, 2011), Lim and Ham (2014)), in which messages are sent by a player to the one who moves before him (e.g., the sender can indicate to the player who moves before him that he will take a certain action). One advantage of using "forward directed" communication is that explanations such as belief-independent lying aversion (see Gneezy (2005) and Vanberg (2008)) by Player C do not apply to our context.[8]

4.2 Experimental Procedure

The experimental procedure in every session is as follows: At the start of each round, subjects were assigned their roles (Player A, B or C) and played the game shown in Figure 1. Each payoff point in the game corresponded to 14 cents in monetary terms. After the subjects made their decisions but before the results for the round were revealed, we elicited the beliefs of subjects. Specifically, we measured τ_{AB}, δ_{BC} and τ_{CAB} from Players A, B, and C respectively.[9]

[8] We note that there are some studies which show that "forward directed" communication may not effectively induce cooperation (see Charness and Dufwenberg (2006) and Andreoni and Rao (2011)), because receivers may not respond well to direct expectations imposed by senders. We do not find evidence for this in our experiments.

[9] In each round we asked each subject only one question about their beliefs to avoid introducing too much complexity into the experiment.

The first order beliefs τ_{AB} and δ_{BC} are not of critical importance, and they provide some useful benchmarks for assessing how Player A and B behave according to their beliefs. Although we did not incentivize the elicitation of these two beliefs, we have no reason to expect that the subjects would report them untruthfully. On the other hand, since τ_{CAB} is the primary concern of the experiment and due to the more complex nature of this belief, we incentivized the elicitation of τ_{CAB}.

Eliciting τ_{AB} provides a basis for us to incentivize the elicitation of τ_{CAB}. In what follows, we briefly explain the procedure of eliciting τ_{AB} first, then go through the elicitation of τ_{CAB} in detail.

Subjects in the role of Player A who had chosen Left were asked to guess the probability (answered in percentage points) that Player B would have chosen Left had he (Player A) chosen Right. Note that in this case, we asked Player A about his beliefs in a hypothetical scenario. For the Player A who had chosen Right, he was asked to guess the probability that the Player B matched in the same group had chosen Left.

The questions posed to subjects in the role of Player C were more intricate. Our goal was to obtain a measure of the indirect higher order belief τ_{CAB}. Player Cs were told that the question that will be asked of them depended on the choice made by the Player A that was matched with them. If Player A had chosen Left, they were asked to guess Player A's answer to the question about Player B's probability of choosing Left (in the hypothetical case that Player A had chosen Right). If Player A had chosen Right, Player C was also asked to guess Player A's answer about the probability that Player B had chosen Left. To motivate the subjects in the role of Player C to think carefully about these questions, they were awarded an extra 25 cents if their answers came within 10 points of the actual guesses of the Player As (which were provided in percentage points) who were in the same group. In the experiment, after reading the instruction to the subjects, we gave verbal quizzes to make sure that the subjects understood all the belief elicitation questions clearly, as well as the fact that only Player C's questions were incentivized.

After subjects completed their guesses, the outcomes of the game for that round were shown on the computer screens. Subjects who were Player C were also told whether their guesses fell within 10 points of the actual guesses of the Player As who were in the same group.

We chose above belief-elicitation protocol because it is simple and easy to describe in the instruction. Our idea is to obtain a rough but meaningful estimate of the subjects' beliefs (particular τ_{CAB}) and we focus on examining the variations in these beliefs across treatments.[10]

[10] A common concern in experimental literature involving belief elicitation is that elicitation of beliefs may induce subjects from non-belief-based thinking towards belief-based thinking and subsequently change their choices (see

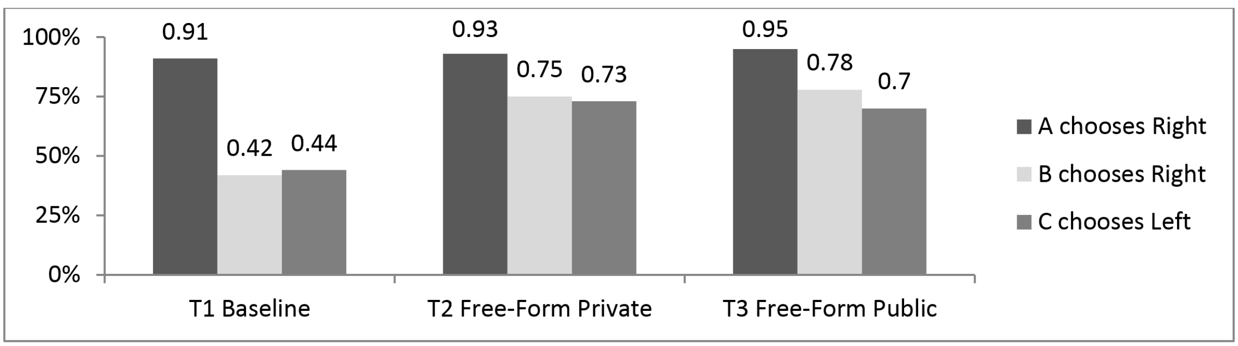
4.3 Results of Baseline and Free-Form Communication Treatments (T1-T3)

4.3.1 Effect of Communication on Choices

Figures 2 and 3 display the proportion of Player A choosing Right, the proportion of Player B choosing Right (given that Player A had chosen Right), the proportion of Player C choosing Left (given that Players A and B had chosen Right) and the distribution of the game's outcomes across the baseline treatment (T1), the *AtoC Free-Form Private Treatment* (T2) and the *AtoC Free-Form Public Treatment* (T3).

In the baseline treatment (T1), 131 of 144 (91%) Player As chose Right. Given this, 55 of 131 (42%) Player Bs chose Right, and this was followed by 24 of 55 (44%) Player Cs choosing Left. The SPE outcome was the modal one, occurring 53% (76 of 144 groups) of the time, while the Cooperation outcome occurred only 17% (24 of 144 groups) of the time.

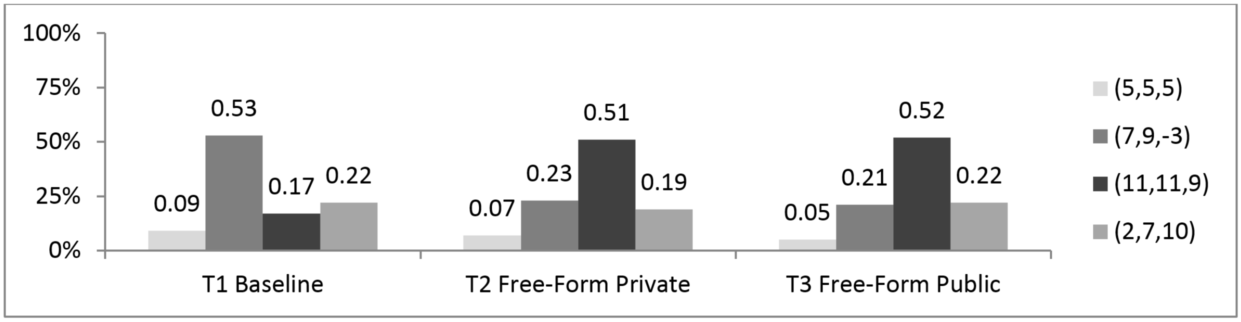
Figure 2: Conditional Frequencies of Choices in T1, T2 and T3



In the *AtoC Free-Form Private Treatment* (T2), 134 of 144 (93%) Player As chose Right, 101 of 134 (75%) Player Bs chose Right conditional on Player A choosing Right, and 74 of 101 (73%) Player Cs chose Left given that Players A and B chose Right. The SPE outcome now occurred only 23% (33 of 144 groups) of the time, while the incidence of the Cooperation outcome was 51% (74 of 144 groups). In the *AtoC Free-Form Public Treatment* (T3), 205 of 216 (95%) Player As chose Right; given this, 160 of 205 (78%) Player Bs chose Right and 112 of 160 (70%) Player Cs chose Left. The SPE outcome occurred 21% (45 of 216 groups) of the time, while the incidence of the Cooperation outcome was 52% (112 of 216 groups).

Rutstrom and Wilcox (2009) for a discussion). Our study assumes that decision makers employ belief-based thinking and investigates how communication influences the beliefs of decision makers.

Figure 3: Proportions of Outcomes in T1, T2 and T3



From these results, one can readily observe that allowing Player A to communicate to Player C reduced the proportion of the SPE prediction of (7, 9, -3) while increasing the proportion of the Cooperation outcome of (11, 11, 9).

We now evaluate the statistical significance of the treatment effects on the players' behavior and the game outcomes. Recall again that each subject played the same role 4 times. We account for this by clustering the observations at the subject level in our tests. Next, for each subject in each role, we use a binary variable to determine whether they did the following: 1) Choose Left more than 50% of time, or 2) choose Left less than or equal to 50% of time. Table 1 compares these decisions across T1, T2 and T3 for Player B and Player C.

Table 1: Test for Effects of Communication on Player B's and C's Choices and Cooperation Outcomes

Frequency of B choosing > 50% L			Frequency of C choosing > 50% L			Cooperation Outcome (11, 11, 9)		
T1	T2	Z Stat	T1	T2	Z Stat	T1	T2	Z Stat
16/35	7/36	2.36***	12/28	27/36	-2.61***	24/144	74/144	-6.22***
(46%)	(19%)		(43%)	(75%)		(17%)	(51%)	
	T3			T3			T3	
	8/54	3.21***		37/53	-2.36***		112/216	-6.75***
	(15%)			(77%)			(52%)	

*** indicates p<0.01, one-tailed tests.

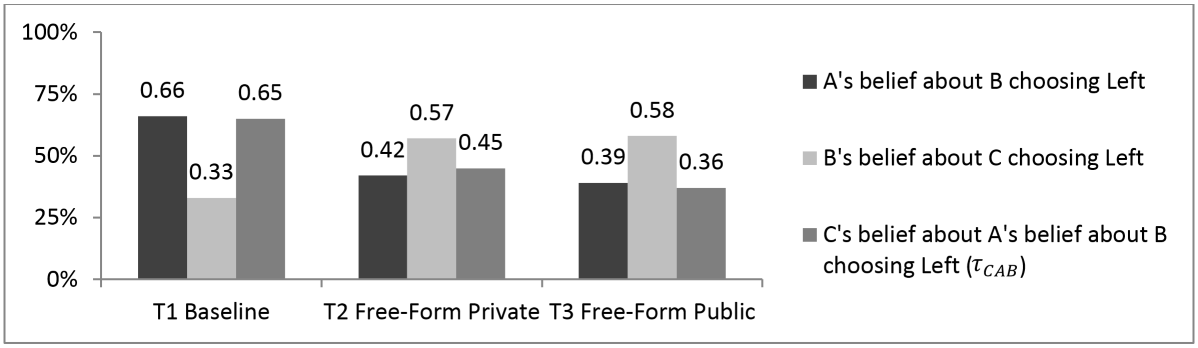
The statistical tests in Table 1 demonstrate that the increase in the proportion of the Cooperation outcome in T2 and T3 compared to T1 is driven by the fact that Player Bs' frequency of choosing Right and Player Cs' frequency of choosing Left increase in T2 and T3, compared to T1.

Note that we find no significant differences in the proportion of choices and the Cooperation outcomes between the T2 and T3 treatments. One possible explanation would be that in T2, even though Player B could not observe the message content sent by Player A to Player C, he knows that Player A could send a message that encouraged Player C to choose Left (and that both Players A and C know this as well).[11]

4.3.2 Messages, Beliefs and Behavior

Figure 4 displays the means of the beliefs that we elicited (τ_{AB}, δ_{BC} and τ_{CAB}) across the three treatments.

Figure 4: Beliefs of Players A, B and C



In T1, the average beliefs were $\tau_{AB} = 0.66$, $\delta_{BC} = 0.33$ and $\tau_{CAB} = 0.65$. In T2, these were $\tau_{AB} = 0.42, \delta_{BC} = 0.57$ and $\tau_{CAB} = 0.45$, while they were $\tau_{AB} = 0.39$, $\delta_{BC} = 0.58$ and $\tau_{CAB} = 0.36$ in T3. One can observe that compared to the baseline T1, τ_{AB} is lower and δ_{BC} is higher, and in particular, τ_{CAB} is lower in T2 and T3 where communication was allowed.

Table 2: Tests for the Effect of Communication on Players' Stated Beliefs

A's belief (τ_{AB})			B's belief (δ_{BC})			C's Belief (τ_{CAB})		
T1	T2	t Stat	T1	T2	t Stat	T1	T2	t Stat
0.66	0.42	4.32***	0.33	0.57	-3.84***	0.65	0.45	3.65***
	T3			T3			T3	
	0.39	5.21***		0.58	-4.48***		0.36	5.66***

*** indicates p<0.01, one-tailed tests.

Next, we relate the beliefs and choices of Player C. From Tables 1 and 2, one can readily observe the relationship between a lower indirect higher order belief τ_{CAB} and a higher incidence of Player C choosing Left, as suggested by our model.

[11] This is likely given that each subject rotated among all three roles across the 12 decision rounds in the experimental session.

To better understand the effect of communication on players' beliefs and behavior, we divide the messages sent by Player A to Player C in T2 and T3 into two categories: "Cooperative" or 'Non-cooperative" messages. "Cooperative" messages include those that explicitly state one's intention to reach the Cooperation outcome, or those that directly request Player C to choose Left. The remaining are classified as "Non-cooperative" messages.[12] Table 3 shows some examples of the two types of messages the participants sent in T2 and T3:

Table 3: Examples of Messages by Player A to Player C in T2 and T3

T2 Free-Form Private	**Examples**
"Cooperative" Messages	*"I'm going to push right so that gives you a chance to make more and me to make more so please choose left"* *"I'm sure you know, choosing left give the highest total payoff. ;)"* *"right right left. all is fair and square"* *"I am going to choose right, hopefully b does too, and you should choose left!"* *"hey! it would be super dandy if you would choose c!"*
"Non-cooperative" Messages	*"how much does a polar bear weigh? enough to break the ice"* *"I'm going left. Sorry about it."* *"I'm coastin on a dream, west coastin"* *"heyyyyeyyyyyeyeyyeyeyyyy"* *"milwaukee brewers"* *"the guy up front has a nice stache ha! I am going right"* *"long live A$AP"*

T3 Free-Form Public	**Examples**
"Cooperative" Messages	*"let's do A right B right C left"* *"RIGHT RIGHT LEFT, we all get paid the most and can buy beer"* *"A-right B-right C-left it is the best way"* *"right, right, left. pay it forward and have faith in humanity"* *"please choose left player C haha"* *"You should choose left so it will maximize almost everybodies profit."* *"C- dont be greedy for the extra 14 cents"*
"Non-cooperative" Messages	*"I don't know what to say, so we can all leave the chat room"* *"Happy Friday, folks"* *"the dude's milio's looks really good"* *"equality for all"* *"let's milk the school for money"* *"Player B go left, C will screw us over if you don't"* *"B you should pick left because C is always greedy"*

[12]We do not further classify the "Non-cooperative" messages due to the small sample sizes for these.

Table 4: Tests for the Effect of Message Type on Cooperation Outcome

	Cooperation Outcome (11,11,9)		
Treatment	"Cooperative" Message	"Non-c" Message	t Stat
T2	55/96	19/48	2.00**
	(57%)	(40%)	
T3	95/163	17/53	3.31***
	(58%)	(32%)	

** and *** indicate p<0.05 and p<0.01, respectively, one-tailed tests.

Table 4 shows that the Cooperation outcome of (11, 11, 9) was reached more frequently when Player A sends a "Cooperative" message, compared to when the message was "Non-cooperative", in each of the two treatments.

Next, we investigate if the type of messages sent by Player A affects the beliefs of players in T2 and T3. Table 5 shows that in T3, the mean τ_{AB} of those Player As who sent "Cooperative" messages is significantly lower than that of those who sent "Non-cooperative" messages (0.36 versus 0.48, $p < 0.01$). Also, the mean τ_{CAB} of Player Cs who observed "Cooperative" messages is significantly lower than that of those who saw "Non-cooperative" messages (0.34 versus 0.44, $p < 0.01$). In addition, the mean δ_{BC} of Player Bs who saw "Cooperative" messages is significantly higher than when they saw "Non-cooperative" messages.

Table 5: Tests for the Effect of Message Type on Players' Stated Beliefs

	A's belief (τ_{AB})			B's belief (δ_{BC})			C's Belief (τ_{CAB})		
Treatment	"Cooperative"	"Non-c"	t Stat	"Cooperative"	"Non-c"	t Stat	"Cooperative"	"Non-c"	t Stat
T2	0.43	0.38	1.23				0.44	0.48	-0.93
T3	0.36	0.48	-3.44***	0.6	0.51	2.28**	0.34	0.44	-2.42**

** and *** indicate p<0.05 and p<0.01, respectively, one-tailed tests.

In T2, however, the differences across the different message types, for τ_{AB} and τ_{CAB} for Player A and Player C, respectively, were not statistically significant. One plausible explanation for this finding is as follows: Because Players A and C know that Player B cannot see the message content in T2, there is no reason for them to expect Player B to change his behavior depending on the message type. Note nevertheless that both τ_{AB} and τ_{CAB} are significantly lower in T2 compared to T1, because Player B still knows that Player A can send a message to Player C.

Overall, the experimental results across T1-T3 confirm our prediction that Player C will choose Left more frequently (leading to a higher incidence of the Cooperation outcome) as his indirect higher order belief τ_{CAB} decreases.[13]

4.4 Results for the Fixed Message Treatment (T4)

We now turn to the results for the Fixed Message Treatment (T4) where Player A can send a predetermined message to Player C. Recall that the message Player A can choose to send was "*I intend to choose Right, because I believe that Player B will choose Right as well.*" Note that Player B does not know if Player A chooses to send this message.[14] Our rationale for this treatment is as follows: By sending the message, Player A is directly communicating to Player C that his reason for choosing Right is because his belief about the probability of B choosing Left, τ_{AB}, is low. If Player C believes Player A's message, then we would expect his stated indirect higher order belief τ_{CAB} to be low as well. Therefore, one would expect the incidence of the Cooperation outcomes to be higher in T4 than in the baseline treatment T1. Furthermore, because the predetermined message directly manipulates the indirect higher belief τ_{CAB}, we also expect δ_{BC} to be higher and τ_{AB} to be lower in T4 compared to the *AtoC Free-Form Private Treatment* (T2).[15]

Table 6: Tests for Differences in Player B's and C's Choices and Cooperation outcomes between T1, T2 and T4

Frequency of B choosing > 50% L			Frequency for C choosing > 50% L			Cooperation Outcome (11, 11, 9)		
T4	T1	Z Stat	T4	T1	Z Stat	T4	T1	Z Stat
1/36	16/35	-3.87***	30/36	12/28	3.38***	106/144	24/144	-9.71***
(3%)	(46%)		(83%)	(43%)		(74%)	(17%)	
	T2			T2			T2	
	7/36	-2.25**		27/36	0.87		74/144	-3.89***
	(19%)			(75%)			(51%)	

** and *** indicates p<0.05 and p<0.01, respectively, one-tailed tests.

In T4, 128 of 144 (89%) of Player As sent the fixed message to Player Cs. 139 of 144 (97%) of Player As chose Right; conditional on this, 128 of 139 (92%) of Player Bs chose Right, and given

[13] As discussed in Charness and Dufwenberg (2006), there may be a reverse causality between a player's direct higher order belief and his behavior due to a false consensus effect (Ross et al. 1977). For example, in the two-player experiment of Charness and Dufwenberg's (2006), Player B might think that other Players B's will choose like him and that Player A's beliefs lean in this direction too (a false consensus effect). Hence, Player B's belief about Player A's belief (about Players B's behavior) merely resemble his own behavior. In our three-player centipede game, such a problem is avoided because τ_{CAB} is not a direct higher order belief. Hence, τ_{CAB} does not reflect Player C's behavior.

[14] We chose to keep the message choice private to Player B to avoid manipulating Player B's beliefs δ_{BC} too strongly (and similarly, so that the effect of δ_{CBC} on Player C's behavior is not too strong).

[15] We compare T4 and T2 because they are both private treatments.

this, 106 of 128 (83%) of Player Cs chose Left. The SPE outcome occurred only 8% (11 of 144 groups) of the time and the incidence of the Cooperation outcome was 74% (106 of 144 groups).

Table 6 shows that the increase in the proportion of Cooperation outcome in T4 compared to T1 and T2 is driven by the fact that player Bs' frequency of choosing Right and Player Cs' frequency of choosing Left increase in T4 compared to T1 and T2.

Table 7: Tests for Differences in Players' Stated Beliefs between T1, T2 and T4

A's belief (τ_{AB})			B's belief (δ_{BC})			C's belief (τ_{CAB})		
T4	T1	t Stat	T4	T1	t Stat	T4	T1	t Stat
0.34	0.66	-6.35***	0.63	0.33	5.06***	0.34	0.65	-6.49***
	T2			T2			T2	
	0.42	-1.58*		0.57	1.11		0.45	-2.07**

*, ** and *** indicates p<0.1, p<0.05 and p<0.01, respectively, one-tailed tests.

Table 7 shows that τ_{CAB} is indeed significantly lower in T4 than in T1 and T2, which is consistent with our prediction. In addition, τ_{AB} is lower in T4 than in T1 and T2. Finally, δ_{BC} is higher in T4 than in T1. Although there was no statistical difference in δ_{BC} between T4 and T2, the directional difference was as predicted.

5 Conclusion

We experimentally examine why communication may matter for inducing cooperation in a three-player centipede game in which the first and the third players do not interact sequentially, but only through the second player. We posit that the third player's decision to help or hurt the first player depends on his indirect higher order belief, that is, his belief about what the first player believes the second player would choose. The evidence demonstrates that communication between the first and the third player can effectively induce cooperation from the third player through shaping his indirect higher order belief.

Indirect higher order beliefs can play an important role in contexts where some players do not interact sequentially, but their actions have consequences on all players. This is especially true in organizations such as bureaucracies and large corporations, where there are many hierarchies or project teams. Hence, we believe that extending our current research to other multiple-player games in which indirect higher order beliefs may matter and testing the role of communication in these games would be a fruitful research agenda. For example, consider a game that starts with

Player A choosing to let either Player B or Player C make the next move. The chosen player then makes a decision that ends the game with payoffs for all three players. In this context, there could be ambiguity of intentions because the chosen player (B or C) might not be sure about why Player A had chosen him over the other player. Hence, the chosen player's belief about Player A's belief about the other player's behavior may affect his decision.

Appendix

A A Model of Sequential Reciprocity

In this Appendix, we investigate how the indirect higher order belief of Player C may affect his behavior assuming that instead of guilt aversion, he is driven by reciprocity concerns. Following the sequential reciprocity model of Dufwenberg and Kirchsteiger (2004), we first define the following functions:

$$\pi_C^{e_A} = 1/2[5 + (-3\tau_{AB} + (1-\tau_{AB})(9\delta_{AC} + 10(1-\delta_{AC})))]; \tag{4}$$

$$\pi_C^{e_B} = 1/2(-3 + (9\delta_{BC} + 10(1-\delta_{BC}))); \tag{5}$$

$$\pi_A^{e_C} = 1/2(11 + 2) = 6.5; \tag{6}$$

$$\pi_B^{e_C} = 1/2(11 + 7) = 9. \tag{7}$$

The function $\pi_j^{e_i}$ can be regarded as a norm for player i that describes the "equitable" payoff for player j given player i's beliefs about the other players' behavior. In other words, player i's kindness towards player j is zero if he believes that player j's material payoff is the average between the lowest and the highest material payoff of j. For Player A, given his belief τ_{AB} and δ_{AC}, he knows that if he chooses Left, Player C gets 5, and if he chooses Right, Player C gets $-3\tau_{AB} + (1-\tau_{AB})(9\delta_{AC} + 10(1-\delta_{AC}))$. Therefore, the average $\pi_C^{e_A}$ serves as a reference point that measures Player A's kindness towards C. For Player B, conditional on Player A choosing Right, given his belief δ_{BC}, he knows that if he chooses Left, Player C gets -3, and if he chooses Right, Player C gets $9\delta_{BC} + 10(1-\delta_{BC})$. Similarly, the average $\pi_C^{e_B}$ serves as a reference point that measures Player B's kindness towards C. For Player C, conditional on Players A and B choosing Right, he knows that if he chooses Left, Players A and B both get 11, and if he chooses Right, Players A and B get 2 and 7, respectively. Therefore, $\pi_A^{e_C}$ and $\pi_B^{e_C}$ are the reference points that measure Player C's kindness towards A and B.

Given the "equitable" payoff functions, we can now define kindness. Since our focus in this paper is on Player C, we define only Player C's kindness and his beliefs about Player A's and B's kindness. Player C's kindness towards Players A and B if he chooses Left is given by

$$k_{CA}(L) = 11 - \pi_A^{e_C} = 4.5; \tag{8}$$

$$k_{CB}(L) = 11 - \pi_B^{e_C} = 2. \tag{9}$$

The above measures how much Player C brings Players A and B if he chooses Left, compared to the reference points. Player C's belief about Player A's kindness towards C if A chooses Right is given by

$$\begin{aligned}
\lambda_{CAC}(R) &= -3\tau_{CAB} + (1-\tau_{CAB})(9\delta_{CAC} + 10(1-\delta_{CAC})) - E_C[\pi_C^{e_A}] \\
&= -3\tau_{CAB} + (1-\tau_{CAB})(9\delta_{CAC} + 10(1-\delta_{CAC})) \\
&\quad -1/2[5 + (-3\tau_{CAB} + (1-\tau_{CAB})(9\delta_{CAC} + 10(1-\delta_{CAC})))] \\
&= 1/2[5 - \delta_{CAC} - \tau_{CAB}(13 - \delta_{CAC})]. \qquad (10)
\end{aligned}$$

The above captures how much Player C believes that Player A brings him if the latter chooses Right, compared to his belief about Player A's reference point. One can see that $\lambda_{CAC}(R)$ is decreasing in τ_{CAB}. In other words, as τ_{CAB} increases, Player C believes that Player A is less kind towards him. Moreover, when τ_{CAB} is sufficiently large, $\lambda_{CAC}(R)$ is negative, which means that Player C believes that Player A is unkind to him.

Player C's belief about Player B's kindness towards him when B chooses Right is given by

$$\begin{aligned}
\lambda_{CBC}(R) &= 9\delta_{CBC} + 10(1-\delta_{CBC}) - E_C[\pi_C^{e_B}] \\
&= 9\delta_{CBC} + 10(1-\delta_{CBC}) - 1/2[-3 + 9\delta_{CBC} + 10(1-\delta_{CBC})] \\
&= 1/2(13 - \delta_{CBC}). \qquad (11)
\end{aligned}$$

The above measures how much Player C believes Player B brings him when B chooses Right, compared to his belief about Player B's reference point. One can see that $\lambda_{CBC}(R)$ is always positive, which implies that Player C always interpret B's choice of Right as a kind action.

Given the definitions of Player C's kindness and his beliefs about Player A's and B's kindness towards him, we assume that Player C experiences additional utility from reciprocation if he chooses Left. This extra utility is given by

$$\begin{aligned}
& Y_{CA} \times k_{CA}(L) \times \lambda_{CAC}(R) + Y_{CB} \times k_{CB}(L) \times \lambda_{CBC}(R) \\
= \ & 2.25 Y_{CA} \times [5 - \delta_{CAC} - \tau_{CAB}(13 - \delta_{CAC})] + Y_{CB} \times (13 - \delta_{CBC}), \qquad (12)
\end{aligned}$$

where Y_{CA} and Y_{CB} are exogenous parameters that capture Player C's sensitivity to reciprocity concerns towards Players A and B, respectively. Note that Player C's utility from reciprocating Player B's action is always positive, while Player C's reciprocity utility towards A can be either positive or negative. When this utility is negative, it means that Player C experiences disutility from helping A when he believes that Player A is unkind to him. (Note that if we include this term into Player C's utility from choosing Right, it can be interpreted as a utility gain from reciprocating Player A when he believes that A is unkind to him.)

In sum, the utilities of Player C from choosing Left and Right are given by

$$U_C(Left) = 9 + 2.25Y_{CA} \times [5 - \delta_{CAC} - \tau_{CAB}(13 - \delta_{CAC})] + Y_{CB} \times (13 - \delta_{CBC}); \quad (13)$$

$$U_C(Right) = 10. \quad (14)$$

Again in this model, the direct higher order belief δ_{CAC} may be inversely related to τ_{CAB}. If Player C believes that Player A believes that Player B is likely to choose Right (i.e., low τ_{CAB}), it is also reasonable to assume that Player C thinks Player A chooses Right because Player A believes that Player C is likely to choose Left (i.e., high δ_{CAC}). Given this, Player C chooses Left if

$$9 + 2.25Y_{CA} \times [5 - \delta_{CAC} - \tau_{CAB}(13 - \delta_{CAC})] + Y_{CB} \times (13 - \delta_{CBC}) > 10. \quad (15)$$

One can see that when τ_{CAB} is low enough (so that δ_{CAC} should be high as well) and δ_{CBC} is high enough, i.e., when Player C believes that both Players A and B expect to reach the Cooperation outcome, Player C would choose Left. Therefore, in the sequential reciprocity model, the indirect higher order belief τ_{CAB} works in the same direction in affecting Player C's behavior as in the guilt aversion model.

B Instructions for the AtoC Free-Form Public Treatment (T3)

Instructions for Experimental Game

1. Introduction

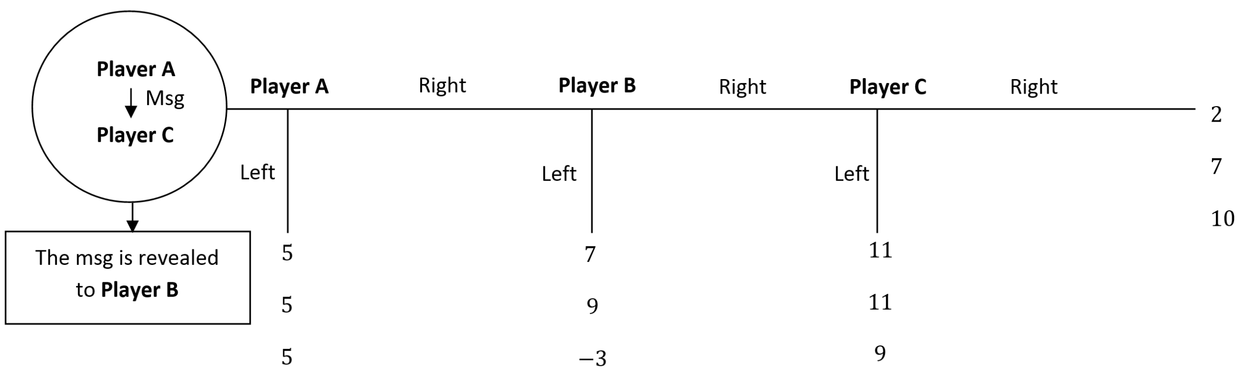
This is an experiment in decision-making. The instructions are simple, if you follow them carefully and make good decisions, you could earn a considerable amount of cash that will be paid to you immediately and privately after the experiment. The amount of cash you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. Do not look at the decisions of others, talk, laugh or engage in any activities unrelated to the experiment. You will be warned if you violate this rule the first time. If you violate this rule twice, we will cancel the experiment immediately and your earnings will be $0.

The participants in this experiment will participate in a total of 12 decision rounds. At the start of each round, you will be assigned to be either Player A, Player B or Player C. The roles will be reassigned each round, which means that in each round you will play a different role. The computer will also randomly and anonymously match the participants into groups such that there is one Player A, one Player B and one Player C in each group. This matching process will be repeated in every round. That is, you will be re-matched with another two players in every round until all 12 rounds are completed.

All participants will play the Experimental Game for the entire 12 rounds. In the Experimental Game, we will always begin with Player A's decision, followed by Player B's decision and then by Player C's decision. In each group, Player A can send a message to Player C at the start of the game. This message will also be revealed to Player B. We will use a computer program to coordinate the experiment. The specific decisions for each player are described below.

2. Moves and Decisions

<u>Experimental Game:</u>



Player A moves first:

- Player A sends a message to Player C. Player A can type whatever she wants to say to Player C. Please type your message in the dialog box in 90 seconds.
- After the message is sent, both Player B and Player C can see the message sent by A.
- Player A chooses either LEFT or RIGHT.
- If Player A chooses LEFT, the game ends and Players A, B and C each receive 5 points.

If Player A chooses RIGHT, then Player B gets to make a decision:

- Player B chooses either LEFT or RIGHT.
- If Player B chooses LEFT, the game ends and Player A receives 7 points, B receives 9 points and C makes a loss of 3 points.

If Player B chooses RIGHT, then Player C gets to make a decision:

- Player C chooses either LEFT or RIGHT.
- If Player C chooses LEFT, Player A receives 11 points, B receives 11 points and C receives 9 points.
- If Player C chooses RIGHT, A receives 2 points, B receives 7 points and C receives 10 points.

PAYOFF TABLE FOR THE EXPERIMENTAL GAME

	A receives	B receives	C receives
A chooses LEFT	5	5	5
A chooses RIGHT, B chooses LEFT	7	9	-3
A chooses RIGHT, B chooses RIGHT, C chooses LEFT	11	11	9
A chooses RIGHT, B chooses RIGHT, C chooses RIGHT	2	7	10

3. Guessing what the other players will choose

After every round, depending on whether your role was that of Player A, Player B or Player C, we ask that you answer the following questions. We ask that you answer the questions according to what you believe to be true.

Question for Player A: The question you answer depends on whether you had chosen LEFT or RIGHT.

*If you had chosen LEFT, please answer the following question:

I guess that if I had chosen RIGHT, there is a ( ) percent chance that Player B would have chosen LEFT. (Fill in an integer from 0 to 100.)

*If you had chosen RIGHT, please answer the following question:

I guess that there is a ( ) percent chance that Player B has chosen LEFT. (Fill in an integer from 0 to 100.)

Question for Player B: You will be asked to answer a question ONLY if Player A has chosen RIGHT. The question you answer also depends on whether you had chosen LEFT or RIGHT.

*If you had chosen LEFT, please answer the following question:

I guess that if I had chosen RIGHT, there is a ( ) percent chance that Player C would have chosen LEFT. (Fill in an integer from 0 to 100.)

*If you had chosen RIGHT, please answer the following question:

I guess that there is a ( ) percent chance that Player C has chosen LEFT. (Fill in an integer from 0 to 100.)

Questions for Player C: The question you answer depends on whether Player A had chosen LEFT or RIGHT. You will receive payment for a correct answer.

*If Player A had chosen LEFT, your question is as follows:

Player A has chosen LEFT and has been asked to guess the probability (in percentage terms) that Player B would have chosen LEFT.

My guess is that Player A's GUESS is ( ) percent. (Fill in an integer from 0 to 100.)

*If Player A had chosen RIGHT, your question is as follows: Player A has chosen RIGHT and has been asked to guess the probability (in percentage terms) that Player B has chosen LEFT.

My guess is that Player A's GUESS is ( ) percent. (Fill in an integer from 0 to 100.)

Note: If your guess (of Player A's guess) is within 10 percent of Player A's actual guess, you will receive 25 cents.

4. Cash Earnings

Your final earnings include the following two components:

1. Player C will be rewarded 25 cents for each correct guess.

2. You will earn also 14 cents $\times$(total payoffs in the 12 rounds). In other words, each point in the Experimental Game is worth 14 cents.

References

Andreoni, J., Rao, J., 2011. The Power of Asking: How Communication Affects Selfishness, Empathy, and Altruism. Journal of Public Economics 95, 513-520.

Attanasi, G., Nagel, R., 2008. Survey of Psychological Games: Theoretical Findings and Experimental Evidence. A. Innocenti and P. Sbriglia (eds.) Games, Rationality and Behaviour. Essays on Behavioural Game Theory and Experiments, Palgrave McMillan, Houndmills and New York, 204-232.

Bartling, B., Engl, F., Weber R., 2014. Does Willful Ignorance Deflect Punishment?-An Experimental Study. European Economic Review 70, 512-524.

Bartling, B., Fehr, E., Herz, H., 2014. The Intrinsic Value of Decision Rights. Econometrica 82, 2005-2039.

Bartling, B., Fischbacher, U., 2012. Shifting the Blame: On Delegation and Responsibility. Review of Economic Studies 79, 67-87.

Battigalli, P., Dufwenberg, M., 2007. Guilt in Games. American Economic Review 97, 170-176.

Battigalli, P., Dufwenberg, M., 2009. Dynamic Psychological Games. Journal of Economic Theory 144, 1-35.

Battigalli P., Charness, G., Dufwenberg, M., 2014. Deception: The Role of Guilt. Journal of Economic Behavior and Organization 93, 227-232.

Baumeister, R., Stillwell, A., Heatherton, T., 1994. Guilt: An Interpersonal Approach. Psychological Bulletin 115, 243-267.

Baumeister, R., Stillwell, A., Heatherton, T., 1995. Personal Narratives About Guilt: Role in Action Control and Interpersonal Relationships. Basic and Applied Social Psychology 17, 173-198.

Berg, J., Dickhaut, J., Mccabe, K., 1995. Trust, Reciprocity, and Social History. Games and Economic Behavior 10, 122-142.

Chang, L., Smith, A., Dufwenberg M., Sanfey, A., 2011. Triangulating the Neural, Psychological, and Economic Bases of Guilt Aversion. Neuron 70, 560-572.

Charness G., Cobo-Reyes, R., Jiménez, N., Lacomba, J., Lagos, F., 2012. The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game. American Economic Review 102, 2358-2379.

Charness G., Dufwenberg, M., 2006. Promises and Partnership. Econometrica 74, 1579-1601.

Charness G., Dufwenberg, M., 2011. Participation. American Economic Review 101, 1211-37.

Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests. Quarterly Journal of Economics 117, 817-869.

Charness, G. Rabin, M., 2005. Expressed Preferences and Behavior in Experimental Games. Games and Economic Behavior 53, 151-169.

Chen, H., Lim, N., 2013. Should Managers Use Team-Based Contests? Management Science 59, 2823-2836.

Coffman, L., 2011. Intermediation Reduces Punishment (and Reward). American Economic Journal: Microeconomics 3: 77-106.

Drugov, M., Hamman, J., Serra, D., 2014. Intermediaries in Corruption: An Experiment. Experimental Economics 17, 78-99.

Dufwenberg, M., 2008. Psychological Games. Entry for the New Palgrave Dictionary of Economics (2nd edition).

Dufwenberg, M., Gächter, S., Hennig-Schmidt, H., 2011. The Framing of Games and the Psychology of Play. Games and Economic Behavior 73, 459-478.

Dufwenberg, M., Gneezy, U., 2000. Measuring Beliefs in an Experimental Lost Wallet Game. Games and Economic Behavior 30, 163-182.

Dufwenberg, M., Kirchsteiger, G., 2004. A Theory of Sequential Reciprocity. Games and Economic Behavior 47, 268-298.

Erat, S., 2013. Avoiding lying: The Case of Delegated Deception. Journal of Economic Behavior and Organization 93, 273-278.

Fehr, E., Herz, H., Wilkening, T., 2013. THe Lure of Authority: Motivation and Incentive Effects of Power. American Economic Review 103, 1325-1359.

Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological Games and Sequential Rationality. Games and Economic Behavior 1, 60-79.

Glaeser, E. G., Laibson, D., Scheinkman, J. A., Soutter, L. C., 2000. Measuring Trust. Quarterly Journal of Economics 65, 811-846.

Gneezy, U., 2005. Deception: The Role of Consequences. American Economic Review 95, 384-394.

Güth, W., Ockenfels P., Wendel, D., 1994. Efficiency by Trust in Fairness? Multiperiod Ultimatum Bargaining Experiments with an Increasing Cake. International Journal of Game Theory 22, 51-73.

Hamman, J., Loewenstein, G., Weber, R., 2010. Self-Interest Through Delegation: An Additional Rationale for the Principle-Agent Relationship. American Economic Review 100, 1826-1846.

Hannan, L., Kagel J., Moser, D., 2002. Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior. Journal of Labor Economics 20, 923-951.

Kolpin, V., 1992. Equilibrium Refinement in Psychological Games. Games and Economic Behavior 4, 218-231.

Lai E., Lim, W., 2012. Authority and Communication in The Laboratory. Games and Economic Behavior 74, 541-560.

Lim, N., Ham, S., 2014. Relationship Organization and Price Delegation: An Experimental Study. Management Science, 60, 586-605.

Oexl, R., Grossman, Z., 2013. Shifting The Blame to A Powerless Intermediary. Experimental Economics 16, 306-312.

Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. American Economic Review 83, 1281-1302.

Ross L., Greene, D., House, P., 1977. The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes. Journal of Experimental Social Psychology 12, 279-301.

Rutstrom, E., Wilcox N., 2009. Stated Beliefs versus Inferred Beliefs: A Methodological Inquiry and Experimental Test. Games and Economic Behavior 67, 616-632.

Vanberg, C., 2008. Why Do People Keep Their Promises? An Experimental Test of Two Explanations. Econometrica 76, 1467-1480.