Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

*Amy Yuan*

# Information Storage and Recall in Indirect Reciprocity

MSc Thesis 2010

# Information Storage and Recall in Indirect Reciprocity

***Name:*** Y. Yuan

***ANR:*** 379399

Supervised by Dr. M. Takezawa

Number of Words: 19,000

Tilburg University

09th March, 2010

# Content

# Abstract

This paper is mainly focusing on the evolution of cooperation, in particular the principle of the indirect reciprocity. And it is based on an experimental cooperation game, which is designed to study indirect reciprocity exclusively and the strategic reputation building up. Our experiment demonstrates that by providing additional second-order information, the cooperation rate does increase. The second-order information offers individuals more possibilities to justify the action and reputation. Players still tend to adopt image scoring or strict discriminator strategy during the game rather than the standing strategy, which might be triggered by the large demand of memory capacity on information storage and recall. The tendency of justifiable cooperation has been performed though due to the fact that the strategy people adopt is more likely to share the definition of goodness in three common features: giving to 'good' as 'good', not giving to 'good' is 'bad', and not giving to 'bad' as 'bad'. Even though people tend to use both first- and second-order information, the accurate information storage in the memory and recall for the second-order information is rather difficult and invisible. Behavior and intuition shown from the post questionnaires indicates consistent results derived from actual action people taken during the experiment – more likely to make information storage and recall when there is only first-order information comparing with the full information condition.

*Key Words: Evolution of Cooperation, Indirect Reciprocity, Information Storage, Memory Recall, Experimental Economics, Behavior Economic, Reputations*

# Introduction

Human beings are featured as possessing the tendency of cooperation even though there might be no expectation of future interaction. In nowadays society, it seems like less and less people are willing to cooperate without receiving much higher benefits as return. However, that does not necessarily indicating the disappearance of indirect reciprocity. The term of indirect reciprocity is initially represented by Alexander in 1987, who argued that persons' behavior towards others is not only affected by their own experience but also by their observations of other's behavior (Engelmann and Fischbacher, 2002: 2). Moreover, according to Alexander, reputation is the key information for individuals to make their decision on cooperation (Engelmann and Fischbacher, 2002: 2). Both social scientists and economists are starting to get interested in indirect reciprocity because one-shot interaction between strangers in the international market is becoming more and more frequent, which induce the tendency towards replacing the long-lasting association and exchanges based on repeated give and take between relatives, neighbors or people from the same social circle (Nowak and Sigmund, 2005: 1291) Many transactions and auctions these days are held no longer face-to-face (i.e. online transaction), such interactions between anonymous partners would raise questions like how the trust is built up, what are the possibility of moral hazard, how effective the reputation is with rather limited information (Nowak and Sigmund, 2005: 1291)? Economists also have noticed the important role played by the reputation for sustaining the cooperation in an institution or within a small group of people that are familiar with one another's history (Bolton et al., 2005: 1458).

In the whole large society, the effectiveness of reputation is much less certain due to the fact that most players are strangers or only knowing about one another through word-of-mouth (Bolton et al., 2005: 1458). While word-of-mouth is not necessarily providing accurate information about the third party but with rather low costs comparing with legal contracts as an example of formal institutional intervention. Thus, several issues about the principle of indirect reciprocity remains as puzzles – what is the drive for each individual to help the others even to the strangers, what might be the factors that influence individual's decision on cooperation, what are the available information for everybody (i.e. reputation), do individuals store relevant information about the others, whether people recall and use such pieces of information for judgments and so on.

Since the breakthrough point proposed by Nowak and Sigmund (1998), more attention has been drawn on the further understanding and developments of indirect reciprocity. When individuals decide to cooperate with the strangers, whom they only might be able to meet maximally once in their life, what are the rules or criterions used by those individuals when they are making the decision of helping the others or not? Questions such like those can be answered by taking a deeper looking at the theories concern indirect reciprocity. Section 1.1 will show the evolution of indirect reciprocity which has been processed so far. In particular, direct reciprocity will be shortly explained first, in order to get clearer image of the emergency of indirect reciprocity. After that, classical models or games that used for explaining/testing

indirect reciprocity such as image scoring and standing strategy will be explained in Section 1.2. And the third sub section of this chapter will discuss all the relevant theories of indirect reciprocity.

## 1.1 Evolution of Indirect Reciprocity

Until the 1990s, majority research on cooperation has developed focusing on restricted (direct) exchange, in which an exchange of altruistic acts happens between the same two actors (Takehashi and Mashima, 2007: 159). A classical example of direct reciprocity is the prisoner's dilemma game, which only have two actors A and B in one cycle using Tit-for-Tat (TFT) strategy – if A helps B, B will be expected to help A next time; whereas if A does not choose to help B, B will then not choose to help A either in the future (see Figure 1). If we setting the benefit value as $b$, and the cost of giving as $c$ (where $c < b$), then for the A helps B, B also helps A back case, both of the agents would have a net gain ($b - c$); while for the A defects from B, B defects from A as well case, no transaction would have occurred – no gain, no loss. However, expectation does not necessarily the same as the reality: if A first helps B, B would receive a net benefit of $b$. Then when B has to choose whether to help A in the next time, B could defect from A. As a result, B will have a payoff of $b$, but A will end up with losing $c$. It implies that the payoff-maximization move is defecting in a one-shot game[1] – a unilateral defector would earn the highest payoff, and the

[1] However, people behave differently if the prisoner's dilemma game is repeated. Repeated game can show sustainable cooperation by trigger strategy if the probability of future rounds is sufficiently high (Nowak and Sigmund, 2005: 1292). Because a rational person should evaluate the benefits of helping

exploited cooperator would not receive any benefits (Nowak and Sigmund, 2005: 1292). In evolutionary game theory, it is not assumed that people are rational, but only that successfully strategies spread – for example, by being inherited, or through learning (Nowak and Sigmund, 2005: 1292). For direct reciprocity, it typically considers a well-mixed population that individuals meet randomly and play sets of Prisoner's Dilemma games with each other, then accounts for the total payoff. When the cooperation occurs, a complex evolution that depends on size of population, the cost-to-benefit ratio, the average number of rounds played, and the probability of errors would happen (Nowak and Sigmund, 2005: 1292). In particular, Axelrod and Hamilton (1981) have proved the TFT as an evolutionary stable strategy (ESS) (Axelrod and Hamilton, 1981: 1393). During the computer round robin tournament for the Prisoner's Dilemma conducted by Axelrod, among all the strategies, TFT strategy which with the features of never be the first to defect, possibility of revenge, and forgiveness after just one act of revenge that scored the highest payoffs (Axelrod and Hamilton, 1981: 1393). It implies that TFT is an ESS for the direct reciprocity (Axelrod and Hamilton, 1981: 1393; Ohtsuki and Iwasa, 2006: 435).

However, the real society is much more complicated than only have two actors playing For example, in the case of the whole population formulates a big social network that contains all the actors in the society, in which case that is much closer to

the other person in the same circle for the current round against the cost of no helping interaction with the other person in the all future rounds, he/she then would choose to cooperate (Nowak and Sigmund, 2005: 1292).

the reality and the daily life, there is no fixed circle that only involves the same actors[2] all the time. All the individuals are in the same pool, they have the right to choose whom they are willing to help with, however, once paired these two persons will never have another interaction any more. In this sense, players are playing one-shot Prisoner's Dilemma game in each round. Such setting ensures the exclusion of direct reciprocity and allows the maximum possibility to investigate in the indirect reciprocity. It implies that for instance, Player A and Player B now are randomly met, and Player A is facing a choice of whether to help this Player B. The fact is that Player A never knows about Player B personally, but Player A might happen to know Player C, whom had interacted with Player B already or heard things about Player B from the others. Or Player A might coincidently observed about the interaction between Player B with the other individuals before. Or there might have existing public information about everybody if information about everybody's movement on cooperation is institutionalized. Player A then has to make a decision based on the rather limited information gathering through 'word-of mouth' or observation about his/her partner in this shot (see Figure 2) if there is no transparent public information about everybody's social/image score. And that is more likely to be the case in the real world simply because formal institutional intervention such as legal contract is way too expensive (Bolton et al., 2005: 1458). It is expected that Player A would choose to help Player B, if Player C tells Player A that Player B has been a good person that helping him/her as well/or the others, or if Player A has observed the cooperation between Player B and

[2] Here, the same actors should be more than 2, otherwise, it is the direct reciprocity

the others by him/herself[3]. This is so-called indirect reciprocity. In this context of indirect reciprocity, two same actors are only supposed to be pair with each other for the interaction at most once, which implies every player can be randomly potentially paired from the whole population, but will only be experiencing give to and receive from with the same other person maximally once (Nowak and Sigmund, 2005: 1292) In other words, Player A's giving to Player B should not be reciprocated by Player B's giving to Player A, but by Player B's giving to Player C or the others (where Player C or the others are not directly involved within this one-shot interaction of Player A and Player B by continuing with the earlier example..Moreover, study done by Nowak and Sigmund (2005: 1292) shows that indirect reciprocity shares a similar model with the one mentioned above for direct reciprocity. It is assumed that strangers meet randomly within a well-mixed population, one (for instance person A) will potentially play the role of a giver, and the other person (for instance person B) will play as a potential receiver. This A and B participate in several rounds of interaction in both roles but only are paired with the same partner once at most. Still, the total payoff is the only determinant. Two strategies are possible for individuals to adopt – an unconditional strategy or a conditional strategy. An unconditional strategy includes choosing always to cooperate and always to defect, while a conditional strategy discriminates the potential receivers based on past interaction (Nowak and Sigmund, 2005: 1292). The discrimination among potential receivers can be seen as a process of building up the reputation. Let us take another similar example, if person 1 helped

[3] Here we only explain the general setting of indirect reciprocity, so no distinguish between helping the 'good' or helping the' bad' yet

person 2, person 3 has observed such an interaction. Person 3 is more likely to consider person 1 as a good person if he/she had been helped by person 2 before. Because person 2 is a good person due to the fact that he/she had chosen to cooperate with person 3 in the past. Now, person 3 sees that person 1 helped person 2, which means person 1 helped a good person. Thus, person 1 is also a good person according to person 3's criteria. In the future, person 3 will also choose to help person 1 if that takes place. Nowak and Sigmund (2005: 1292) pointed that if the cost-to-benefit ratio is sufficiently low, and the amount of the information about the co-player's past is sufficiently high, cooperation based on discrimination can emerge. In the resulting population, cooperation is more likely to occur towards to the individuals who have helped (Nowak and Sigmund, 2005: 1293). However, Nowak and Sigmund (2005: 1293) also figured that two questions concerning such an indirect reciprocity model explained above would appear: one is that why should individuals care about the other individuals' action rather than their own payoffs? The other issue is the lack of stability of the cooperation. Wonders like that need to be solved by finding the possible ESSs for the indirect reciprocity, which comes in subsequently in the next subsection.

### 1.2 Image Scoring or Standing Strategy

In order to answer the wonders left from the model of indirect reciprocity, it is important to first develop the possible strategies people adopt. Most commonly mentioned strategy chosen in the literatures are image scoring strategy and standing

strategy, Image scoring strategy was initially proposed by Nowak and Sigmund in 1998, they showed that the strategy of helping those who have helped others could evolve to certain fix part of the population (Milinski et al., 2005: 2495). A simple scoring system was used as the basis of simulating the results. At the beginning of each game, all the players in the population were assumed to have an image score of zero. Rule for accumulating the image scoring is that if a player helped the other person, his/her image score would be increased, otherwise his/her image scores would be decreased by withholding the help to that other person (Milinski et al., 2001: 2495). By applying such an image scoring system to the indirect reciprocity model in subsection 1.1 the strategies for helping would range from, for instance, -5 to +6 (Milinski et al., 2001: 2495). An image score of -5 means the individual is almost unconditionally helping, while an image score of +6 indicates the participant is almost unconditionally not helping because he/she only cooperates with the people with an image score of at least +6 (Milinski et al., 2001: 2495). Such a strategy that scoring increases with helping and decreasing with defecting is so called image scoring strategy.

Image scoring strategy however has been shown unstable by ignorance of the other possible existing strategies the others could be adopting at the same time. And individuals who adopt image scoring strategy sometimes can hurt some innocents. Individuals who use the image scoring strategy only would choose to help the person whom helped his/her partner in the previous round, and would not be willing to help a

‘bad’ person whom did not help his/her partner in the last rounds (Takahashi and Mashima, 2006: 419). As a result, the image score of the individuals who is employing image scoring strategy would still decrease by not helping such a ‘bad’ person, and leading to further non-cooperation due to low level of image scores. It turns out that people are not actually acting according to their own interest on helping the others, but rather due to the fact that only the action of help (no matter a ‘good’ or a ’bad’ person) increments the image score. The rational player thus would choose to ignore the image score of potential recipients, and try to increase his/her own score (Milinski et al., 2001: 2496). Leimar and Hammerstein (2001) concluded that the evolution of cooperative image scoring strategy only happens when there are restrictive conditions that include either a substantial impact genetic drift or a very small cost of helping (Milinski et al., 2001: 2496). That implies that the image scoring strategy uses only first-order information (Takahashi and Mashima, 2007: 162). The individuals who adopt such strategy calculate the score of the other individuals by assuming everyone is good at the very beginning, then treat the ones who gave to another individual when being assigned as a giver the last time as ‘good’ and treat the ones who did not help the other individual when being assigned as a giver in the previous round as ‘bad’. Even though the way of taking image scoring strategy is rather simple, subsequent studies have shown it cannot maintain generalized exchange because sometimes people who use image scoring strategy hurt one another (Takahashi and Mashima, 2007: 162). As a result, the ones who adopts image scoring strategy would have less expected payoffs than the ones who adopts All-C strategy

(unconditionally cooperate), because with All-C strategy people can never be treated as a 'bad' (Takahashi and Mashima, 2007: 162). And the proportion of All-C would increase until the population is too vulnerable to convert into All-D (unconditionally defect), which eventually leads to the entire collapse of the indirect reciprocity (Takahashi and Mashima, 2007: 162)

After three years' of proposal on image scoring strategy by Nowak and Sigmund (1998), by following Sugden's strategy which is targeting at 'good standing' defined in 1986, the standing strategy has been proposed by Leimar and Hammerstein (2001) and Panchanathan and Boyd (2003) (Takahashi and Mashima, 2006: 419). The image score system remains, but in a rather different accounting method. In Sugden's model (1986), everybody is in good standing at the very beginning – a person would fail to keep good standing if he/she does not help a potential recipient in good standing, whereas withholding the help to the ones whom lack of good standing does not decrease the image scores. The main difference between image scoring strategy and standing strategy thus is that for an individual it is his/her own interest to react to the standing of a potential recipient. In the other words, not cooperate with the people with bad reputation does not decrease individuals' own scores. The standing strategy overcomes the weakness from the image scoring strategy, because it distinguishes justifiable from unjustifiable defection, and for the ones who adopt the standing strategy, they do not lose the opportunities to receive benefits from others who also using the standing strategy (Takahashi and Mashima, 2007: 163). It implies that the

standing strategy is more sustainable for indirect reciprocity than the image scoring strategy. And things are also getting more complicated due to the introduction of second-order information together with the standing strategy (Takahashi and Mashima, 2007: 163). As a result, the standing strategy shows the stability on the evolution of cooperation and plays the game with more fairness.

### 1.3 Extending Strategy Space – Leading Eight

As showed in subsection 1.2, to those who have 'good' social reputation does a player give aid as reciprocation, whereas he/she has to refuse to help those who have 'bad' reputation; otherwise, benefits of altruism is easily exploited by them (Ohtsuki and Iwasa, 2006: 107). The remaining puzzle for all the strategies thus is about how to define the 'goodness' of the reputation. Ohtsuki and Iwasa (2006: 107) then examined condition for ESSs over 4096 possible cases with the combinations of reputation dynamics and behavioral strategies exhaustively, and mathematically proved that only eight cases so-called 'leading eight' are crucial to the evolution of indirect reciprocity. There are three commonly recognized features for the 'leading eight': (1) giving to the individuals with 'good' reputation should be considered as a 'good', (2) not-giving to the individuals with 'good' reputation should be considered as a 'bad', (3) not-giving to the individuals with 'bad' reputation should be considered as a 'good'. (Takahashi and Mashima, 2006: 420)

On the other hand, as the same setting earlier, one donor and one recipient, and the

donor assigns a score to a recipient by using first-order information (the recipient's previous action) and second-order information (the reputation score of the current recipient's previous recipient) that is built up on the basis of the recipient's previous behavior as a donor (Takahashi and Mashima, 2007: 161). Study done by Takahashi and Mashima (2007: 161) shows the proper way to represent all the possible strategies that have been mentioned in the literature framework is by the sets of four genes[4]: Gene 1 determines whether a current recipient who helped his/her own recipient whose score was 'good' when he/she was in the position of a donor is evaluated as 'good' or 'bad'; Gene 2 determines whether a current recipient who helped his/her own recipient whose score was 'bad' when he/she was in the position of a donor is evaluated as 'good' or 'bad'; Gene 3 determines whether a current recipient who did not help his/her own recipient whose score was 'good' when he/she was in the position of a donor is evaluated as 'good' or 'bad'; Gene 4 determines whether a current recipient who did not help his/her own recipient whose score was 'bad' when he/she was in the position of a donor is evaluated as 'good' or 'bad' (See Table 1). As a result, GGBB is representing image scoring strategy, and GGBG is representing standing strategy.

However, study done by Takahashi and Mashima (2007) shows that the strict discriminator strategy (SDISC) ('GBBB') is the solution for making indirect

[4] Each gene determines whether a particular type of other should be considered 'good' or 'bad', and all strategies give to a 'good' recipient and do not give to a 'bad' recipient (Takahashi and Mashima, 2007: 161).

reciprocity possible, when there is second-order information presenting. Individuals who adopt SDISC strategy assigns 'good' only to the people who gave to a 'good' receiver (Takahashi and Mashima, 2007: 163). There is a distinct contrast between the standing strategy and SDISC strategy: the standing strategy distinguishes justifiable from unjustifiable defection, while SDISC strategy distinguishes justifiable from unjustifiable cooperation (Takahashi and Mashima, 2007: 163). As a result, the standing strategy treats giving to 'bad' as equal to giving to 'good', whereas SDISC strategy punishes those who give indiscriminately (Takahashi and Mashima, 2007: 163). But according to the mentioned three commonly features of the 'leading eight''s, the standing strategy is with much more consistency than SDISC is (Takahashi and Mashima, 2006: 420). And SDISC is actually not included in the 'leading eight', but many strategies in the line with the 'leading eight' commonly shared features failed to maintain indirect reciprocity in Takahashi and Mashima (2003)'s simulation (Takahashi and Mashima, 2006: 420). Nevertheless, Takehashi and Mashima (2005b, 2006) conducted selective play, in which individuals are actively select their desirable recipient if they know the scores of all the individuals, and there is not enforcement for individuals to interact with randomly matched recipients (Takahashi and Mashima, 2007: 164). Results show that the fourth gene does matter little in the selective play, simply because it is always possible for a donor to find at least one desirable recipient from the population in the selective play, which would imply that SDISC and 'GBBG' are equally successful (Takahashi and Mashima, 2007: 165). While extra-standing strategy ('GBBG') is indeed in the line with common criteria on the 'leading eight',

and was included as one of the strategies that is ESS. Thus, in the selective play environment, SDISC could also be considered as an ESS strategy.

Moreover, possible errors in perceptions can be presented, which could be subjective, simply because the perception of the other people as 'good' or 'bad' differs from individuals, even if those individuals might employ the same strategy; which could also be objective when misperception is calculated once for the whole population (Takahashi and Mashima, 2007: 164). It is argued that whether an error in perception is shared by the whole population could have serious potential implication (Takahashi and Mashima, 2007: 164): assume misperception is subjective, at one point, individual 1 with the standing strategy and is regarded as 'good' by the others. Individual 1 treats individual X as a 'bad' person by mistaken, while the others who are using the same standing strategy as individual 1 treat individual X accurately as a 'good' person. In the later on, when individual 1 is paired with X, he/she would choose to defect from X because of his/her own perception. And this might proceed further because misperception is subjective. Consequently, the ones who adopt the same standing strategy might hurt one another. But All-C individuals ('GGGG') do not get involved because perception is irrelevant to them, which might lead to All-C individuals to evolve and All-D individuals ('BBBB') take over the population in the end (Takahashi and Mashima, 2007: 164). However, such a circle does not take place if the perception is objective, in which case each individual's score is shared among all the individuals (Takahashi and Mashima, 2007: 164). Ohtsuki and Iwasa (2004)'s

study shows that by assuming perception error is objective, the 'leading eight' could maintain the indirect reciprocity in (Takahashi and Mashima, 2007: 164). As well in the case of SDISC, the perception error is subjective or objective does not matter that much because it treats All-C individuals as 'bad' by definition (Takahashi and Mashima, 2007: 164).

To sum up, there is one disagreement between Ohtsuki and Iwasa (2004, 2006) and Takahashi and Mashima (2006, 2007) – whether the definition of goodness should include individuals who give to 'bad'. Ohtsuki and Iwasa (2004, 2006) argue that it should include individuals who give to 'bad' when define goodness because it is more important to distinguish justifiable from unjustifiable defection and to assign 'good' to justifiable defection. In contrast, Takahashi and Mashima (2006) argue that it should not include individuals who give to 'bad' when define goodness since the indirect reciprocity can be only maintained if the individuals who give to 'bad' are punished in some ways (Takahashi and Mashima, 2007: 165, 171). But they do share some main points in common: the ESS strategies for the indirect reciprocity must use both first-order and second-order information, and give to 'good' is the crucial feature of all the strategies that account for the emergence and maintenance of the indirect reciprocity (Takahashi and Mashima, 2007: 172) . In this sense, even though each strategy uses different criteria to determine who to help, indiscriminate altruism (i.e. All-C strategy) does not exist (Takahashi and Mashima, 2007: 172).

On the other hand, after explaining the 'leading eight', now we will move on to see whether there are other literatures supporting the same conclusion found by Ohtsuki and Iwasa (2004, 2006) and Takahashi and Mashima (2006, 2007).. According to Bolton et al. (2005), the efficient outcome of the helping game requires all the givers to play give, in spite of the incentives to play keep. As a result, the available information must be sufficient to provide a long-term relationship for rewarding the cooperation and punishing the defection. The asymmetry of the helping game excludes the existence of direct reciprocity, allowing a concentrated focus on the indirect reciprocity (Bolton et al., 2005: 1459). When there is no information about the reputation, the cooperation from indirect reciprocity is more like to be unsustainable, unless the reputation information is becoming transparent (Bolton et al., 2005: 1459). Theoretically, the need of recursive information on reputation should be enormous, which results the invisibility of the stable indirect reciprocity except for the special situations (Bolton et al., 2005: 1459).

Recursive information is mostly gathered by backward induction. The equilibrium can be obtained from backward induction if the stop time for the process is indefinite and with complete information, but the experiment we conduct is with a finite, scheduled, and under monitoring (Bolton et al., 2005: 1459). Obviously, the assumption for the cooperation in indirect reciprocity can be derived from backward induction is way too strong. Therefore, our model assumes players have very limited ability to perform backward induction (Bolton et al., 2005: 1459). Then discrimination becomes to be an

evolutionary stable strategy in a helping game with finite horizon, however, discriminating the others solely based on the first-order information is very short-sighted, which might even cause the collapse of the cooperation in the end (Bolton et al., 2005: 1460). Therefore, adding a second-information would sustain/stabilize the cooperation in the indirect reciprocity (Bolton et al., 2005: 1460). Therefore, Bolton et al. (2005) also reached the same conclusion – both first- and second-order information should be used when making a decision of helping whom from a different perspective.

In conclusion, image scoring strategy and the standing strategy are the most commonly mentioned strategies individuals adopt when they are playing such a helping game. But the image scoring strategy only uses the first-order information which causes some weakness such as hurting innocents. In order to overcome those weak points, the standing strategy was introduced together with the second-order information that might help to sustain the cooperation in the indirect reciprocity. However, the subjective or objective misperception may cause totally different scenarios for the people who employ the standing strategy. Takahashi and Mashima (2007) shows the solution to properly use the second-order information and do not confuse between subjective and objective misperception, SDISC strategy can sustain the cooperation in the helping game designed for indirect reciprocity, even though such strategy was not included in Ohtuski and Iwasa (2004)'s 'leading eight'. More importantly, Ohtsuki and Iwasa (2004, 2006), Takahashi and Mashima (2006, 2007),

and Bolton et al. (2005) all show that any evolving strategies for the indirect reciprocity should be using both first- and second-order information.

### 1.4 Experimental Investigations of Indirect Reciprocity

There had been several past experiments conducted on investigating if people do use both first- and second-order information among strangers. Following we will discuss three past experiments on the indirect reciprocity.

Milinski et al. (2001) had run an experiment for distinguishing between image scoring and standing strategy. They assigned one person to be 'NO player', whom plays as a defector all the time. Results show the players do take notice of the second-order information. However, the strategy people adopt in the experiment is more in the line with the image scoring strategy or a similar one to a larger extent (Milinski et al., 2001: 2495). It might because the standing strategy requires too much working memory capacity (Milinski et al., 2001: 2495). Moreover, donors of the NO player compensated for their refusing to help these players by being more generous to others (Milinski et al., 2001: 2495). Such compensation could not due to the adoption of the standing strategy, but may have helped the donors of the NO player to achieve a reasonable pay-off, the loser of the game were the players who always received No from the NO player (Milinski et al., 2001: 2500).

In the experiment conducted by Bolton et al. (2005), the classical model of subgame

perfection implies that enforcing cooperation requires recursive information on reputation in nature: one player needs to know not only one's partner's past action but also one's partner's partners' past actions, etc (Bolton et al., 2005: 1458). Within a social network fulfilled with strangers, and lack of a formal tracking institution, information about each other thus is rather less in detail that leads to a relatively harder process to build up reputation. The model with relaxing subgame perfection requirement, however, suggests that sustainable cooperation might still take place based on information such as only about a partner's recent past that is relatively easy to recall (Bolton et al., 2005: 1458). Results show that the probability of cooperation is higher when the second-order information is available together with the first-order information. It implies that the additional second-order information does increase the probability of cooperation[5], thus information has effect in the experiment. Moreover, the tendency to give is strongly influenced by the first-order information about how the receiver treated others in the past; there is a remarkable difference between the first- and second-order information condition: while the probability of giving to a receiver with a 'keep-history' are almost identical across all three information conditions, second-order information appears to provoke much more giving to receiver with 'give-history' than first-order information does (Bolton et al., 2005: 1464). Overall, the increase in the probability of giving when the second-order information is provided is more than twice amount than the increase in probability with only the first-order information (Bolton et al., 2005: 1467). It then implies that

[5] We only would mention the results concerning to the High-cost treatment from Bolton et al., because that is the setting relevant to ours.

candidates are using the second-order (recursive) information to stabilize the indirect reciprocity, and also suggests that when the second-order information is provided, more likely a receiver's action would have the opportunity to be justified and promotes higher probability of cooperation comparing with when only first-order information is available (Bolton et al., 2005: 1467).

Furthermore, according to Takahashi and Mashima (2007), Gene 1 and Gene 2 were evaluated differently, which would appear that people do consider second-order information; a person who gave to 'bad' was evaluated as generous but was evaluated negatively on many other aspects; respondents were less willing to give their resources to a person who gave to 'bad' than they were to give to a person who gave to 'good'. It is consistent with the theoretical argument that generalized exchange can only be maintained if individuals who give to 'bad' are punished in some way, which is also a proof of the social dilemma research that discovered punishing those who do not punish free riders is necessary to achieve mutual cooperation.

In all the past experiments on the indirect reciprocity, experimenter tends to offer at least one scenario with both the first- and second-order information available for the participants. Even though those experiments have found evidence that people do take notice and tend to use both pieces of information, whether players do still use the second-order information when they need to store the information in the memory and retrieve them if necessary remains vagueness.

The second-order information on individual's reputation can be gained either by directly information from the other players, public scoring or recall from player's own memory system where the reputation relevant pieces are stored. Past experiments are mainly offering the second-information through direct information from others or public scoring. It shows that information is spontaneously obtained as a form of reputation from the others, which can be defined as 'reputation-based indirect reciprocity'.

On the other hand, when the second-order information is not directly provided, people then need to make information storage on the past about how everybody has been acting through observations or 'word-of-mouth'. Afterwards, occasionally if person A has to make a choice of helping against person B whom he/she has never interacted before, person A would recall from his/her memory in order to see person B's reputation. It expected to result cooperation if person A consider person B as a good person in his/her memory system. Such an indirect reciprocity can be defined as 'memory-based indirect reciprocity'. That is where our research question emerges, thus our experiment is focus on the circumstance that when information are obtained solely from the memory system whether people do recall and use the memory-based pieces if facing the cooperation decision to make. As majority of the theoretical and empirical studies of indirect reciprocity have pointed out, it is crucial to not only monitoring partners' action/reputation but also all individuals' action/reputation

within the same social network (Nowak and Sigmund, 2005: 1291). It implies that indirect reciprocity requires information storage and transfer as well as strategic thinking (Nowak and Sigmund, 2005: 1291). As a result, our experiment is designed to see the strategies, cases of actions individuals take under different information scenarios. In each session, players are paired randomly: one donor, one recipient. The donor faces the choice of whether to take a costly action to help the recipient. Even though the cooperation in our experiment is socially efficient, the only way to control the possible existence of free-riding behavior is through the reputation that is built up from accumulating a player's past action (Bolton et al., 2005: 1458). In order to ensure the core of investigation is on the indirect reciprocity, we make sure that pair is formed randomly and does only stay at top for once during the whole progress. The reputation mechanism we use is system of indirect reciprocity as well, in which actions taken within one group of partners are reciprocated by strangers or the third party[6].

This paper focuses on indirect reciprocity principle based on a laboratory experiment of a helping game with three different information scenarios: the first one is with full information (FI) condition (second-order information/recursive information); the second one is with partial information (PI) condition (first-order information/immediate past action); and the third one is with no information (NI). Such a laboratory experiment is in terms of the cooperation among strangers when

[6] Such a reputation mechanism is quite similar to the ones in the study done by Bolton et al. (2005)

there is limited information available manipulating in different scenarios. Our main focus is on investigating whether individuals do storage information about others' past behavior, and recall from memory when such piece of information is needed. The second chapter will talk about the design of the experiment in more details.

# Design of the Experiment

So far, most of the theoretical framework in the area of direct or/and indirect reciprocity is derived or tested from mathematical analysis or computer simulation. Here, in order to investigate the memory-based indirect reciprocity, we offer individuals different information condition scenarios, when conducting a computerized helping game. Our experiment is modified based on Bolton et al. (2005)'s study. The experiment was conducted at the ESP laboratory of the Faculty of Social and Behavioral Sciences (FSW) from Tilburg University., which was programmed and carried out with the software named z-Tree (Fischbacher, 2007).

## 2.1 Helping Game

The helping game we organized is involving with information storage and recall in the context of indirect reciprocity. In each round, each player needs to make choices against all the other players within the same group of 7 people about whether to help them if he/she is an assigned the role of a giver. Then 3 pairs are formulated by the

experimenter and roles of giver and receiver were giver. The payoff of the game would be based on the decisions made by each individual at the beginning of the current round. In more details, each giver would have 1 euro in their hands. It implies that in each decision making time, 7 players have to image that he/she has 1 euro, and has to decide whether to give this one euro to the others. If he/she decides to keep this one euro, in this round, he/she would end up with 1 euro in his/her account if he/she was assigned as a giver; while if he/she makes the choice to give it to the others, then he/she will end up with zero account still if assigned as a giver. On the other hand, if the participant was assigned as a receiver, who got helped from the others, he/she will earn 1.6 euro in his/her account; otherwise, he/she would stay with the zero account. It implies that $b$ = 1.6 euro (benefit), and $c$ = 1 euro (cost), and cash is the only exchange, incentive for the players.

In each condition, participants played 3 sessions each consists of 7 rounds of the game. We instructed that one session is randomly selected for determining the final payment. A minimum 3 euro will be paid as a participation fees[7], and maximally 7.8 euro could be earned by one individual. Post questionnaires were collected at the end of the experiment.

[7] In this case, if a player who ends up with an earning that is less than 3 euro, he/she still will be paid by 3 euro.

## 2.2 Information Feedback

Such a helping game manipulates the available information/feedbacks under three different conditions (see I. in Appendix): the 1st condition (FI) provides Full Information including both first- and second-order information to the candidates. During the session public ID number was fixed. Furthermore, ID numbers of 3 givers, ID number of 3 receivers and actions made by the givers were publicly informed at the end of each round. Thus, in this condition, participants could track a history of actions of all the players (i.e. who helped whom at which round). But, those pieces of information were not displayed when participants make a decision at the beginning of the round. If they want to use the information, they need to recall them from their own memory. It implies that players can observe other players actions and they could remember the feedbacks and take such piece of information into account when next time they have to make a decision on whether helping the others That is an application of storing and using the second information – recall from the memory about who helped whom, who had been helped. Moreover, the payoff for each round and accumulated for each session will show at the feedback screen as well when the information about the actions is displayed.

The second condition (PI) only offers Partial Information to the candidates (see I. in Appendix), which is different from the first condition only in one point. ID numbers of receivers were not given at the end of the round. In this condition, the second-order information was not available as it was impossible to know whether a give helped a

good person or bad person. It is an example of using first-order information because only giver IDs and the actions givers chosen are shown on the feedback screen.

In the third condition (NI) (see I. in Appendix), no information including ID numbers were given. Thus, this condition is identical with a series of one-shot helping game. The only piece each candidate would know after each round finishes is their own earnings.

Order of the conditions was fixed for all the groups we conducted: starting with FI, followed by PI and then NI condition.

This way of controlling the available information/feedbacks to the participants during the game allows us to investigate in questions such as whether players do use the first-order and the second-order information if it is available, whether individuals do information storage and retrieval from their memory about the first- and/or second-order information when it is necessary.

### 2.3 Other Details

We excluded the possibility of direct reciprocity so that a pair of players interacts only once and all the participants know about at most once pairing in advance when they read the instructions handed out by the experimenter. For instance, seven players for one group: player 1, player 2, player 3, player 4, player 5, player 6, player 7. If in the

first session of FI, player 1 was assigned as a giver, and player 2 was paired with player 1 as a receiver. Then in the later on session, player 1 will never pair with player 2 anymore. As a result, the reputation of individuals is based on third party's action. It also tries to minimize the chance of direct reciprocity. Furthermore, identities are anonymous, only the experimenter could monitor IDs, and decisions. No discussions are permitted in the experiment room. On each participant's screen, he/she has to make a choice of whether to help the other six people in the stand point of a giver.

For each condition, three sessions are generalized, and in each session seven rounds are played. It implies that in total for one group, there will be 21 rounds for each condition. In each round, three players will be assigned as a giver, and three other players will be paired with the givers as a receiver. The selection of the giver IDs and receiver IDs are all computerized, also randomly. Moreover, under the FI and the PI, the giver ID would stay the same for every seven rounds, which means under the first and second condition, computer randomly assign the giver IDs and the receiver IDs for each session. However, with the NI, the giver IDs and the receiver IDs are randomized for each round.

**2.4 Participants**

In our experiment, we invited in total 91 students from the Faculty of Social and Behavioral Sciences (FSW) of Tilburg University, and most of them are undergraduate first-year students. Those students are divided into thirteen groups,

which would make seven players in each group fulfill the game. Every group plays the game separately upon the time schedule, and the all seven players are the first time meeting each other – strangers. As mentioned earlier, each candidate needs to click on the computer screen to indicate whether he/.she would like to help the other 6 candidates if he/she is in a position of a giver. In each round, three players will be assigned as a giver, and three other players will be paired with the givers as a receiver.

Next chapter, we will start to illustrate the estimated results by running such an experiment described in this chapter.

# Results

### 3.1 Cooperation Rate

Figure 6 shows the trend of cooperation rate in each condition. Remember that each condition consists of 3 sessions and every new session starts at 1$^{st}$, 8$^{th}$, and 15$^{th}$ round. It is very obvious that the FI condition generates the highest probability of cooperation, and probability of cooperation tends to be lower in the last rounds of each session. Moreover, for the NI condition, quite a flat curve on the probability of cooperation, which indicates there is substantial cooperation take place even there is no information available. That implies that not all the cooperation attributable to image scoring strategy (Bolton et al., 2005: 1462). In general, the PI condition has a

lower level of probability of cooperation comparing with the FI condition but with a rather similar shape. It means that players intend to be more cooperative when second-information is available, which might be only because people believe with full information they should cooperate. It is an illusion in most of individuals' mind, but actually they are not necessary to do so. And when second-order information is available, whether the players actually do storage and use it or not remains as a weak point. Figure 5 compares the overall average of the same variable in each condition. In order to get the overall average, we calculate the average of the probability of cooperation under the FI condition of 21 rounds, under the PI condition of 21 rounds, and under the NI condition of 21 rounds. As a result, three new aggregate level of average on probability of cooperation will be given. Apparently, the FI condition has the highest level of probability of cooperation (50%), the PI condition is with a slightly lower level (42%), and only 23% of cooperation under the NI condition. It shows that by adding information the probability of cooperation does increase.

We conducted repeated 3 (condition) × 3 (session) × 7 (round) ANOVA on the probability of cooperation in each group. There are five effects were statistically significant.

Clearly that all three factors – condition ($F$ (1, 12) =64.67, $p$ = .000), session ($F$ (1, 12) =22.73, $p$ = .000), and round ($F$ (1, 12) =52.77, $p$ = .000) have a very significant effect on the probability of cooperation. It leads to the result that condition, session or

round has a strong impact on the tendency of individuals for choosing to cooperate or not. It is more likely that if condition, session or round differs; the probability of cooperation would also differ correspondently. Moreover, interaction Condition × Session × Round does have a significant level of influence on the probability of cooperation ($F$ (1, 12) = 8.65, $p$ = .012), which refers to that the effect of condition and round is moderated by the session. For instance, in the NI condition, cooperate rate was rather flat in the 2nd and 3rd session but exhibited a decreasing trend in the 1st session. On the other hand, decreasing trend was consistently observed in all the sessions in both the condition with FI and PI. However, differences between these two conditions did not exist in the first 3 rounds of the first session. Interaction effect of Condition × Round also reached a significant level of $p$ value ($F$ (1, 12) = 15.34, $p$ = .002). It implies that the decreasing trend was observed only in the FI and PI condition but not in NI condition.

In a nutshell, the FI condition generates the highest probability of cooperation comparing with the PI and the NI condition, even when there is no information available (under the NI). The pattern of the probability of cooperation in the NI condition is fluctuating around 20% that implies not all the cooperation is featuring as the image scoring strategy states. Moreover, by providing additional information the probability of cooperation increases. And condition, session, or round has very significant effect on the probability of cooperation separately, so do these three factors altogether. While the main possible source for the significant influence from

Condition × Session × Round is the interaction effect of Condition × Round. It suggests that the effect of condition and round is moderated by session, and both the FI and the PI condition perform a decreasing trend but not in the NI condition.

### 3.2 Usage of Second-order Information

We measured the probability of cooperation for four different situations for each participant; CG, CB, DG and DB. All candidates are assumed to be good person at the beginning. Generally speaking, two characters were checked: first character (C or D) – action of a target person at the last time when he/she was a giver ((n–1)$^{th}$ or older round); second character (G, B) – reputation assigned to a receiver of a target person when the latter was a giver the last time. If we want to classify the cases for player B as an example: first, we checked the action of player B when he/she was the giver at the last time (which may be (n-1)$^{th}$ round or earlier). We then checked a receiver in that specific interaction (i.e. when player B was the giver the last time). Imagine that player B was a giver at (n-1)$^{th}$ round and player D was a receiver. We then checked a decision made by a participant on player D at (n-2)$^{th}$ round (i.e. whether or not a participant decided to help player D at (n-2)$^{th}$ round[8]. This action is assumed to be a reputation a participant assigned to the receiver. It implies, for instance, if a participant responded to cooperate to a receiver at the last round, we assumed that he thinks that a receiver is a good person. This is because, in the model of indirect reciprocity, it is assumed that people help a person whom they think 'good'. Therefore,

[8] when receiver was self, we removed that case from the following analyses

CG refers to a situation where a participant observed that a giver cooperated (C) a receiver with good reputation (G) at the last round ($t$ – 1). CB refers to the case that a participant observed that a giver cooperated (C) a receiver with bad reputation (B) at the last round ($t$ – 1). DG refers to the situation that a participant observed that a giver defected (D) a receiver with good reputation (G) at the last round ($t$ – 1). DB refers to the case that a participant observed that a giver defected (D) a receiver with bad reputation (B) at the last round ($t$ – 1).

At the beginning of each round, we asked participants to make a decision whether to help other participants with a strategy method and we calculated the probability of cooperation for each four different case in the following way. First, we did count the number of cooperation for each case for seven candidates from every group in each session under three different conditions, as well the total number of each case for each seven individuals from every group in all nine sessions with three different conditions. Then following the same example earlier, imagine that a participant decided to help player B at the $n^{th}$ round. The number of cooperation in CG case is incremented. To derive the probability of cooperation for each player for each case of all thirteen groups per session per condition (e.g. the number of cooperation in CG case/total amount of CG cases).

### *3.2.1. Effect of Action, Reputation in Each Condition*

Here, we combined[9] the three sessions under the same condition for the same case into one new variable, by taking the average of the three dependent variables in the first step. For instance, the probability for the cooperation of the CG case under the 1st condition is now the new variable, which is the mean of the probability of cooperation in the CG case in the 1st session, the 2nd session and the 3rd session under the FI condition.

Table 2 shows the mean of the probability of cooperation together with the standard deviation and the number of observations under each condition. For the FI condition (see Table 2 (a)), the means in the case of CG and CB are slightly higher comparing with the ones in DG and DB case, which is more likely to show as 'GGBB' strategy – image scoring strategy. Under the PI condition (see Table 2 (b)), 'GBBB' is more likely to be the case, which is strict discriminator (SDISC) strategy. For the NI condition (see Table 2 (c)), all the means are quite the similar level, but at a much lower level comparing to the first two conditions. If using the same criteria[10] as for the other two conditions, All-D is likely to be the strategy players adopt.

---

[9] The reason we are doing such a step is to have a look at the different cases under the same condition. And in the 3.2.3., the detailed results of trend will be illustrated – parts of cases no significant impact from the session, which means collapsing sessions per condition for the new variables is reasonable.

[10] If setting that around or above 0.35 is good, below 0.35 is bad

For measuring the statistical significance, we used repeated 2 (action) × 2 (reputation) ANOVA on the probability of cooperation in all 4 cases for each condition. Under the 1st condition (FI), strong significance ($F$ (1, 84) = 32.98, $p$ = .000) has been found from the action on the case choice, and roughly reaching the significant level ($F$ (1, 84) = 2.78, $p$ = .099) of impact from the reputation, the dependent variables here are the probability of cooperation in all four cases under the FI condition. It implies that people do tend to use both first- and second-order information when they are facing the cooperation decision to make. With the 2nd condition (PI), same strong significant level of impact from the action has been observed ($F$ (1, 83) = 20.34, $p$ = .000), together with a much closer to the significance sourcing from the reputation ($F$ (1, 83) = 3.44, $p$ = .067). Moreover interaction of Action × Reputation does influence the cooperation choice at a significant level as well ($F$ (1, 83) = 4.40, $p$ = .039). The dependent variables here are then the probability of cooperation in all fours cases under the PI condition. When only first-order information is available, people seems to remember more clearly about how they have been treated in the past and take that into account when making helping decisions. For the 3rd condition (NI), none of the sources[11] has shown significance, while the dependent variables are the probability of cooperation in CG, CB, DG, DB case under the NI condition.

### *3.2.2. Effect of Action, Reputation, and Session*

The second step we have taken is to run the repeated 2 (action) × 2 (reputation) × 3

[11] Sources include Action, Reputation, and Action × Reputation.

(session) ANOVA on the probability of cooperation in all 4 cases again but by separating both 3 sessions and 3 condition. We are testing the probability of cooperation in each case in each session under each condition with the action, reputation and session as independent variables.

Figure 4 shows that the probability of cooperation in CG and CB case are higher than the probability of cooperation in DG and DB case in the FI and PI condition.

With the 1st condition (FI), the action still has a strong impact on dependent variables ($F$ (1, 37) = 14.01, $p$ = .001), but the reputation now turns out to be insignificant. And the session has a roughly significant influence on the dependent variables ($F$ (1, 37) = 3.44, $p$ = .072), which corresponds to the shape of the probability of cooperation in each case (see Figure 4 (a)) – when session differs, the probability of cooperation in CB, DG, and DB also differ; only the probability of cooperation in CG is rather flat across sessions. As we can see that under the FI condition CG has a rather flat pattern on probability of cooperation, while CB starts with a higher level of probability of cooperation but became lower from the 2nd session. DG and DB both have a lower level of probability of cooperation comparing with CG and CB. And both have a declining shape. But DG ends up with an even lower level of probability of cooperation in the 3rd session than DB. Under the 2nd condition (PI), the action shows a marginal significance level of effect ($F$ (1, 32) = 3.58, $p$ = .068) on the case choices, but not from the reputation, session or interaction effects. As we can see in Figure 4

(b), the curves for all four cases are rather flat across the session, which can be captured by the insignificance level of the session. It looks like CG has a decrease at the 2nd session, but still ends up the highest level of probability of cooperation among 4 cases. CB and DB both have a decreasing trend, while DG has a slightly increasing pattern. In the end of the PI, the probability of cooperation is ranked from CG, CB, DG then DB, which is in the line of the significant level of impact from action. Then the 3rd condition (NI), none of the significant level of impact has been found from the action, the reputation, the session or the interactions. From Figure 4 (c) though, we see that CB and DB starts at the highest level of probability of cooperation, and average all 4 cases have a decreasing pattern even though CG increases from the 1st session to the 2nd session and DB increases from the 2nd session to the 3rd session.

*3.2.3. Effect of Trends on the Reputation Assignment*

In this subsection, we checked trend effect. We ran repeated ANOVA on the probability of cooperation in each four case in each condition with session as an independent variable. Figure 3 (a) shows the average probability of cooperation in four situations for the FI. We found the probability of cooperation for 4 cases varies between sessions under the same condition – under the FI, only in DG case session showed a very significant impact on the probability of cooperation ($F(1, 82) = 9.50$, $p = .003$), and in CB case a roughly statistical significance has been resulted ($F(1, 53) = 3.40$, $p = .071$). And both cases have a declining trend overall speaking. In particular, the probability of cooperation in DG case is with a largely decreasing trend across

sessions under the 1st condition (0.43 → 0.35 → 0.30), while the means for the probability of cooperation in CB case has decreased dramatically from the 1st session to the 2nd session, but slightly increased for the 3rd session. Still the 3rd session does have a lower mean than the 1st session has. The possible explanation for the bit rise from the 2nd session to the 3rd session could be because the participants are aware of the number of sessions they are going to play by reading the instructions. By reaching the end session of each condition, people intend to choose defect (ending-effect). For the other two cases (CG and DB), session does not result a significant effect on the probability of cooperation, but according to the figure, it is more likely that they both having a declining trend, the probability of cooperation in CG case has decreased from 0.45 in the 1st session to 0.39 in the 3rd session, while in the case of DB, the means of the probability of cooperation has decreased a lot as well from the 1st session to the 3rd session (0.32 → 0.25).

For the 2nd condition (PI), two statistical significances sourcing from session has been observed for the probability of cooperation: one is in CG ($F$ (1, 83) = 6.03, $p$ = .016), and the other one is DG case ($F$ (1, 78) = 3.12, $p$ = .081). From Figure 3 (b), we can see that the trend of the CG case is roughly declining, while for the DG case first increased a lot but decreased from the 2nd session moving to the 3rd session. In particular, the mean of the probability of cooperation in DG case for the 3rd session is even lower than the 1st session here. The possible reason behind it could be also due to the end session approaches. For the left two cases, no significance has been

estimated for the CB and DB case, even though it looks like the means of the probability of cooperation in these two cases perform a decreasing trend according to Figure 3 (b).

Then for the 3rd condition (NI), no significance has been observed. From Figure 3 (c), we only can roughly speaking that trend looks more like that the means of the probability of cooperation in CG, CB, and DG case perform decreasing trends, meanwhile the means of the probability of cooperation in DB case has increased at first from the 1st session to the 2nd session, but drops a bit more, from the 2nd session to the 3rd session, to a level which is slightly lower than the mean of the 1st session.

Table 3 indicates that under the 1st condition (FI), CG case has the highest probability of cooperation (36%) comparing to the other 3 cases. The order of the probability of cooperation for all 4 cases is CG, CB, DG, and DB for all 3 conditions. And the probability of cooperation for the cases under the FI condition is higher than the PI and NI condition, while the probability of cooperation for all 4 cases under the PI condition is higher than the NI condition. Furthermore, givers distinguish whether the receiver's last partner was a cooperator or a defector. In particular, givers are more likely to give to cooperated ones when the receivers' cooperating was 'justifiable': the probability of cooperation for CG case is 42% in contrast to 35% for CB case. However, givers are less likely to give to defectors even when defecting was 'justifiable': the probability of cooperation for DG case is 34% in contrast to 27% for

DB case. It suggests that by offering second-order information, people are more likely to use such opportunity to justify a receiver's action, and promoting for the higher probability of cooperation than when only first-order information presents.

In conclusion, we can see that direct reciprocity exist all the time even though we designed to exclude it for our experiment. The strategies players are more likely to adopt for sustaining the indirect reciprocity both in combining or separate sessions are image scoring strategy, All-D strategy, standing strategy and SDISC strategy. It implies mostly CG and CB have been considered as 'good'. Under the FI, action always shows a significant level of impact on the probability of cooperation in cases, while reputation only shows significant influence when collapsing sessions. The situation is similar under the PI condition as in the FI condition. For NI condition, none of the factors show significance on influencing the probability of cooperation in CG, CB, DG, and DB case. The moving trends of each case are in generally a declining pattern, and CG is usually having the highest probability of cooperation except for the NI condition. Furthermore, people seems like using the second-order information when it is available (under the FI condition), however, people tend to remember more accurate information about the past when there is only first-order information.

### 3.3 Behavior and Intuition

After have seen how the probability of cooperation is performing in general (Section 3.1 and 3.2), it is obvious that the FI condition is having on average a much higher level of probability of cooperation, and followed by the PI condition then the NI condition. The rough picture we can observe therefore is that individuals are more likely to cooperate when both first- and second-order information is presented. However, it remains a wonder on whether people is using the information from recalling the memory or is just when the second-order information is available people intend to cooperate more actively. Thus, post questionnaire is aiming at gaining a clearer picture about the behavior and intuition of the candidates, through which we can see whether the scales the players chosen are corresponding to what they actually did behave during the helping game. Moreover, we also can see what are the possible key points that driven the results earlier. Now we do take into account the way people claim to behave[12]. In order to take analysis on the relationship between cooperation decision and individual's intuition on earnings, as well the impact from information storage/recall on the helping decision, we enter the extra data collected through the 5-scale post questionnaire (see II. in Appendix) for each player in every group on the same sheet that containing the data of the probability of cooperate in CG, CB, DG, DB cases for each session under each condition. Due to the fact that the post questionnaire is only separating parts by different information condition, we take the average of the probability of cooperation in CG case in the 1st, the 2nd and the 3rd session under the

[12] The main source of this is the post questionnaire.

FI condition to collapse the sessions, which creates a new variable so-called the probability of cooperation in CG case under the FI condition. The same steps have done for the CG case under the PI and NI condition, as well for the other three cases in each condition. And those 12 new variables are the dependent variables that will be used in the following analysis to see whether earnings' intuition or information storage/recall has influence on the probability of cooperation for 4 keep or give cases.

We analyze several questions from the post questionnaire – as general questions for all three information condition, we did ask the candidates four questions concerning earnings: first of all, how serious were they interested in earning the money during the experiment to see the seriousness on payoffs for each player; secondly, did they try to maximize only their own earning to see if self-interest on earnings matters; thirdly, did they try to maximize the sum earning of the whole group to see if the overall welfare is important to each person; lastly, did they try to maximize the difference between their own earnings and the average earnings of the others to see if earning competition is important. These general questions are targeting at the intuitions candidates have when they are playing the game. Then we asked them whether they did try to remember as much as they can when seeing the results of decisions made by three givers at the feedback screen for the first and second three sessions to check the level of information storage; as well when making decisions, did they care the past behavior of the others in the first and second sessions to check the potential behavior change, after that they had also been asked did they care how the other people treated

them in the past when making decisions in the first and second three sessions to check the attention people paid on the past treatments. Furthermore, all the candidates were asked during the first three sessions with the full information condition, when they saw the feedback information about giver gave (did not give) the money to the others, did they consider whether or not receiver had (not) given the money to the others when those receivers were selected as giver in the past. Such question can reveal the behavior of the first- and/or second-order information storage and recall.

### *3.3.1. Group Effect*

As a beginning step, we just take a look at the post questionnaires solely by the null hypothesis as there are no significant differences between the groups' mean scale for all four general questions, in which case we used a one-way ANOVA treating mean level of answering scale for all four general questions as our dependent variable, and group as the independent variable. From the means level, it looks like the first general question on seriousness of earning scored the highest mean level (3.74), and the last general question on the importance of earning competition has the lowest mean level (2.25). Results show that none of the effects between groups for all the four general questions have shown significance. It suggests the null hypothesis is not rejected, which implies that overall speaking there are not obvious differences between groups' mean in answers for the earnings. All the groups claimed to be behaved in the similar way. In general, they all evaluated the earning competition between self and the others within the same group is rather less important, and they are all serious in earning the

money, especially it looks like that candidates from around half of the groups care about the payoffs above the average level.

By ending the process for only looking at the post questionnaires for the group effect, we want to know whether groupID has an influence on the answer about whether people care about the past treatment from the others under different information condition, thus, we take a one-way ANOVA with mean level of answering scale for all four general questions as our dependent variable, and group as the independent variable. The null hypothesis is formulated as there are no significant differences between the groups' mean scale for the question on do you care about the past treatment from the others when making decision. The means for the first two conditions are both around 3, and the mean for the NI condition is much lower (mean for caring of the past treatment in the last three sessions = 1.95), which indicates that under the NI condition people are less likely to care about the past treatment from the others. However, that is more likely due to the fact that there is no information available for people to observe how all the players treat each others. The mean level on the caring of the past treatment in the FI condition ranks as the highest (3.34). For the significance, only in the first sessions a significant level of impact from groupID has been observed, which implies that groupID only do significantly influence whether players care about the past treatment from the others under the FI condition ($F$ (12, 89) = 1.99, $p$ = .036), but not under the PI and the NI condition. We also further checked with the repeated 3 (Condition) ANOVA for the same question on

caring the past treatment from the others. GroupID has been treated as the covariance. Results is quite consistent with one-way ANOVA, condition has shown a very significant level of *p*-value ($F$ (1, 88) = 9.44, $p$ = .003). It implies the condition is the main source of causing the different mean on answer of level on caring about the past treatment from the others.

*3.3.2. Seriousness on Earning Money*

Now, we would love to know whether the seriousness of earning money is the motivation for making choices of cooperation under all three conditions. At first, we did run a repeated 2 (action) × 2 (reputation) × 3 (condition) ANOVA. It implies that the dependent variables are the probability of cooperation in each case under each information condition; the independent variable is the first general question – how serious were they interested in earning the money during the experiment. Result shows that condition is the only variable has significant *p*-value ($F$ (1, 68) = 6.35, $p$ = .014), which means that different conditions have influence on what case individuals chose. However, no significance has been found on interactions like Condition × Seriousness of Earning Money, Action × Condition × Seriousness of Earning Money, Reputation × Condition × Seriousness of Earning Money, Reputation × Seriousness of Earning Money, Action × Seriousness of Earning Money, Action × Reputation × Condition × Seriousness of Earning Money, which indicates that when Seriousness of Earning Money interacts with action, reputation and condition, it does not show any significant impact on the cooperate decision. In other words, we can say

that by controlling action, reputation and condition, the seriousness on earning money does not influence the decision of cooperation or the choice of cases players choose in all three information conditions.

However, we have been only testing the whole population for three conditions altogether. It is difficult to see are there any correspondences of the post questionnaires and the behavior during the game, for instance, whether people who is more serious on earning the payoffs for certain condition would behave different from the ones who claim to be less seriousness on earnings. Following that, we did exclude the factor of condition, in order to see whether certain condition might have impact on cooperate decision from seriousness for earning money interacting with action and/or reputation. Three separate repeated 2 (action) × 2 (reputation) ANOVA have been run for each condition taking the first question for general as covariant. Result for the FI condition shows interactions of Action × Seriousness of Earning Money, Reputation × Seriousness of Earning Money, Action × Reputation × Seriousness of Earning Money all show insignificance. It means that seriousness for earning money interacts with action is a poor indicator of whether people choose to cooperate or not, and interacts with reputation is a poor predictor for the change pattern on dependent variables as well, so is the one when seriousness for earning money interacts with both action and reputation under the FI condition. However, we only looked at the whole population together, we cannot easily conclude if there is certain consistency with between how the respondents circled the post questionnaires and the behavior

showed in the experiment, we thus test whether the candidates whom claim to be caring lots about earning the payoffs are behave differently from the ones who show less care about earning money under the FI condition. We separate the population half to half[13], the cutting point chosen here is 4 – half[14] is the ones who selected the scale below 4, another half[15] circled scale 4 or 5. Thus, ones with an answer below scale 4 are the ones who claim to be less care about earning money, whereas ones with scale 4 or 5 are the ones seriously concern about the earnings. Results show that the ones care less about the earnings has a roughly significance on the interaction effect of Action × Seriousness of Earning Money ($F$ (1, 21) = 3.36, $p$ = .081), but not on the other interaction effects like Reputation × Seriousness of Earning Money, Action × Reputation × Seriousness of Earning Money. It means that under the FI condition the action people take is moderated from those less seriousness on earning money intuition, what do have an influence on the choice of cooperation candidates take, but not the reputation. On the other hand, for the ones who seriously care about the payoffs, they intend to take into account of the reputation and try to remember as much as information as they could, as well retrieve from the memory when making decision of cooperation or not. Because reputation itself ($F$ (1, 60) = 4.61, $p$ = .025) and the interaction effect of Reputation × Seriousness of Earning Money ($F$ (1, 60) = 5.30, $p$ = .025) do both have a significant level of $p$-value. Therefore, when degree of

[13] We try to divide the population 50-50. However, it is only a 5-scale questionnaire, which is very hard to achieve 50-50. The cutting points we chosen are by averaging the answers, and take the rounding.

[14] This half accounts 27 out of 91

[15] This another half accounts 64 out of 91

seriousness of earning money differs, the moderator is also different on the probability of cooperation in CG, CB, DG, DB case – action (first-order information) for the players with less seriousness on earning the money, or reputation (second-order information) for the people who care lots on the payoffs under the 1st condition, which actually shows the consistency of the intuition and behavior.

For the PI condition, reputation itself does have an influence on the cooperate decision ($F$ (1, 82) = 6.92, $p$ = .010), and seriousness of earning money interacts with reputation also does have a significant impact on the choice of cases ($F$ (1, 82) = 5.13, $p$ = .026) for the whole population. It implies that under the PI condition with only first-order information, players are actually do try to take into account of the reputation and do potentially make better information storage when the feedback screen shows then recall when the later on rounds come up. No significance of interaction Action × Reputation × Seriousness of Earning Money and Action × Seriousness of Earning Money, which means that even though the players are trying to build up reputation about the others by observing and storing the available information in the memory system, the action – current receiver's behavior towards the previous receiver remains on the similar pattern when the decision about cooperation is about to make.

For the NI condition, no interactions are significant, that is mainly because there is no information available for people to observe others' action or reputation. The

seriousness on earning the payoffs is not showing any impact on the cooperation decisions either.

### *3.3.3. Maximization of Wealth*

Besides the seriousness of earning the money, three other general questions on maximization on wealth have also been asked – did you try to maximize only your earnings; did you try to maximize the sum earnings of all the people; did you try to maximize the difference between your earning and the average earnings of the others. Three repeated 2 (action) × 2 (reputation) × 3 (condition) ANOVA have been conducted to check whether those maximization incentives interact with action and/or reputation, and condition have significant impact on the decision of helping the others or not. The first maximization is about how much people intend to maximize self-earnings (the second general question), only action ($F$ (1, 68) = 6.13, $p$ = .016), condition ($F$ (1, 68) = 5.05, $p$ = .028), and Action × Reputation ($F$ (1, 68) = 7.31, $p$ = .009) are showing significance, but incentives to maximize self-earnings interacting with action, reputation, condition separately or in pair or altogether do not show any significant influence on the helping decision. For the maximization of the whole population's wealth, interaction of Reputation× Maximization of Whole Wealth has a $p$-value that is quite close to significant level ( $F$(1, 68) = 3.27, $p$ = .075). Roughly speaking, reputation together with the incentives on maximizing the whole population's wealth has some impact on the decisions of cases candidates would choose. For the level on maximizing the difference between self-earnings and the

earnings of the others, no significant interaction effects have been found. It means that the degree on caring earnings competition between his/her own payoffs and the others' payoffs interacting with action, reputation, condition altogether or in pair or separately have no impact on the cooperate decision. It means that people do not really compete with each other on the earnings when facing a cooperation dilemma.

### *3.3.4. Second-information Storage and Recall*

After testing the general questions on earnings for all three conditions in common, we now are going to have a look at whether the information has impact on whether choose of cooperation for each condition separately.

We first used the paired-sample T test[16] to see whether players do information storage in their own memory differently under different conditions, whether players pay different attention on the behavior changes when conditions differ, whether players care about the past treatment from the others differently under the first two information conditions (the FI and the PI condition). The null hypothesis here assumes the mean of two paired samples equal for each pair. Results show that the means for all three pairs are all around 3, which is not that comparable because all of them are at the similar level. Only the third pair shows a significant level ($t$ (89, 90) = 2.07, $p$ = .021) that implies the null hypothesis is rejected. It means that the mean on the question about the level players do care about the past treatment from the others is

[16] The paired-sample T test is only testing for the FI and the PI condition.

different when the conditions differ (the FI or the PI condition). While means of the question about the level candidates care about trying to remember as much as information at the feedback screen, and did they care about the past behavior of the others when making decision are equal under the FI and PI condition. However, that significance on the mean difference could be resulting from the big role played by the direct reciprocity.

The second check is about the question did you try to remember as much as you can when seeing the results of decisions made by three givers. We test whether players do information storage in their own memory when they are looking at the feedback screen, and whether such behavior differentiates the decision on cooperation to make in each condition. Result under the FI condition shows that only Action × Information Storage in the first three sessions has a very significant impact on the helping decision ($F$ (1, 82) = 22.65, $p$ = .000). It means that action and information storage interactively influence the choice of cooperation, which implies that players tend to remember as much as they can on the first-order information about the givers' past action in the FI condition. It suggests that people are only trying to remember as much as they can on the first-order information even though the second-order information is also available.

For the PI condition, a very strong effect from Action × Reputation × Information Storage on the helping decision has been observed ($F$ (1, 81) = 9.17, $p$ = .003).

However, it is not obvious to see if the people who claim to make more information storage in their own memory make decisions on helping dilemma differently from the people who indicate had less information storage. As a result, we did separate the population using the cutting point[17] of 3: half[18] of the population includes the candidates who claim to have a scale equal or above 3 (high level of information storage in their own memory), the other half[19] of the population is the people who circled a scale 1 or 2 (low level of information storage in their own memory). We observed that under the PI condition there is a roughly significant level from Action × Information Storage on the helping decision for the people who indicate they have a high level of information storage in their own memory ($F$ (1, 53) = 3.50, $p$ = .067), but none of the interactions have shown significance for the population in which players circle a low level of information storage in their own memory. It implies that people with higher level of information storage in their own memory are the main source for causing a significant level of interaction Action × Reputation × Information Storage under the PI condition. And for the candidates who do information storage in their own memory more are the ones more likely to choose to cooperate.

The other thing we have checked is the usage/recall of second-order information, which is only available for the FI condition. Results show a significant impact on the decisions of cooperation players make from interaction effect Action × Recall

[17] Similar to earlier situation, we try to divide the population 50-50, but due to only 5 numbers for the answers, it is not easy to be exact 50-50

[18] This half has 58 out of 91 candidates

[19] This other half contains 32 candidates

Second-order Information ($F$ (1, 82) = 3.85, $p$ = .053), which means that action together with the usage of second-order information do influence players' decision on helping the others or not. But no other interactions are showing significant level of influence. It implies more attention was paid at the first-order information. However, we cannot conclude that whether the people who claimed to be using the second-order information are the ones more intend to cooperate. In order to see if such tendency exists, we separate the population into two[20]. The part that claiming to have less usage of second-order information are the ones choose below scale 3, who actually show a roughly significant level of interaction effect from Action × Recall Second-order Information in the first three sessions on the cooperate decision ($F$ (1, 34) = 3.31, $p$ = .078). It shows that the ones with lower level of recall on the second-order information paid more attention on the first-order information. The other part of the population with high level of usage of the second-order information is rather resulting non-significance on any interactions. It implies that the people who do claim to use higher level of second-order information do not behave the same when they are making choices on cooperation.

In sum, no difference between groups on the general questions like the seriousness of earning money, maximization of self-earning, maximization of whole group's earnings, importance of the earning competition; whereas there is a significant effect

[20] Try to divide 50-50; however, 5-scale questionnaire is not easy for achieving it. The cutting point we chosen here is scale 3. One part contains 38 players out of 91, who is having answer with scale below 3, the other part contains 52, who claim to be having the scale ranged equal or higher than 3.

from group on the question about whether players care about the past treatment from the others under the FI condition. And condition is the only significant source on the different answers on such question about caring the past treatment from the others. Players with more serious consideration on earning the money tend to pay more attention and use the second-order information, whereas the ones with less incentive on earning the money tend to notice and use only the first-order information when both information is available (the FI condition). For the wealth maximization, people seems to acting as the ones who care about the whole society's benefits/welfare rather than themselves' or compete with the others. For the information storage/recall, we find that when information condition differs, the level of caring on other's past behavior differs. Moreover, even though we offer them the second-order information to check the receivers' reputation, players just tend to remember as much as they can on the first-order information about the givers' action. And the people who do claim to recall the second-order information more often are not behaving the same during the experiment.

After discussing the results, the turn for checking the robustness and consistency of the estimated results explained earlier will come in the following chapter.

## Robustness and Consistency

This chapter will compare the estimated results in chapter 3 with the available

relevant studies to see whether our results are robust and consistent with the past literatures.

We will start the comparison for the results of the overall cooperation rate. Our experiment indicates that by providing the additional second-order information, the cooperation rate has risen from 42% to 50% (from the PI to the FI condition). And when there is no information available, there is still quite some cooperation occurred since the cooperation rate in the NI condition is 23%. Such result is very consistent with the result for the high cost scenario in Bolton et al. (2005)'s paper.

The trend for the cooperation rate is looking very similar to the one in Bolton et al. (2005)'s study, the willingness to give tends to be lower in the last rounds, which implies that the future horizon does influence giving rate; there is substantial giving even in the no information condition that demonstrates immediately not all of the giving is attributable to image scoring.

Moreover, our experiment shows that the CG and CB case tend to have a higher rate of cooperation than the DB and DG case when both first- and second-order information is available. Such results are in the line with the experiment done by Bolton et al. (2005), which suggests that candidates are using the second-order (recursive) information to stabilize the indirect reciprocity, and when the second-order information is provided, more likely a receiver's action would have the opportunity to

be justified and promotes higher probability of cooperation comparing with when only first-order information is available (Bolton et al., 2005).

Generally speaking, the rate of cooperation in CG and CB case are different from each other under the FI condition, which implies that people do use the second-order information. Takahashi and Mashima (2007) also showed the similar results.

Moreover, our experiment shows similar results that have been obtained independently by Bolton et al. (2002) (Milinski et al., 2001: 2500). In the treatment with full information condition (FI), they found that facing a cooperative receiver (with a history of CG or CB meaning the receiver helped a cooperator or helped a defector respectively) increased, and facing a selfish receiver (with a history of DG or DB) decreased the probability that the givers to help, which is in the line with the image scoring strategy. Moreover, DB is justified comparing with DG, there situations were treated marginally differently that is a hint for approaching a standing strategy (Milinski et al., 2001).

Results are also consistent with both backward induction and action discrimination models, reputation information, the cost-to-benefit ratio for cooperation, and the length of the game horizon all play important roles in the decision of strangers to cooperate (Bolton et al., 2005: 1465). First-order information increases the probability of cooperation comparing to no information condition, whereas additional

second-information is more effective than no or only first-order information. However, Bolton et al. (2005) also suspect that such results might be captured by an incomplete information model that supposes some people are more likely to cooperate than others. Partial information about reputation then possibly acts as a noisy signal of a player's type. In particular, incomplete information model predicts cooperation can arise even in finite horizon games, which might also start to explain the importance of a player's own history in his/her decision to cooperate because one's own history is potentially informative of the proportion of cooperative types in the population (Bolton et al., 2005). We however further the analysis by taking the look at the post questionnaire, which shows that people's behavior and intuition in certain degree could influence the probability of cooperation in each case under different information condition.

The estimated results of ours show that CG and CB are usually generating a higher level of probability of cooperation than DG and DB do. It implies that people intend to give/help the cooperators rather than the defectors. However, the evidence on CB and DG cases leading to lower probability of cooperation is not as obvious as expected. From the Figure 4, when the second-information is available (under the FI condition), the probability of cooperation for CB cases ranked as the 2$^{nd}$ place among the 4 cases, followed by DG in the 2$^{nd}$ session, but followed by DB in the 3$^{rd}$ session. It is unambiguous that whether people consider DG better than DB case. It is more likely to be in the line with the image scoring strategy that evaluates both DG and DB case as 'bad', but it is not that obvious that people actually considering DG as a 'bad'

strategy. On the other hand, when there is only first-order information in public, the trends for each case are rather smooth, which suggests that the population is more likely to store the first-order information only when second-order information is not available.

Last but not the least, our results also revealed the tendency of moving towards to the standing strategy. We observed that people do actually make information storage in their own memory as they have claimed in the post questionnaires, significances have been observed in the PI condition. It implies that when the feedback is showing, players do try to remember as much as they can. Milinski et al. (2001) also pointed two hints indicating the possible standing strategy adoption: (1) when full information condition, the rate of receiving defect for the person who always acts as a defector increased faster during the first rounds than where only partial information is available; (2) when full information, a trend that donors of the person who always acts as a defector received fewer defections than were expected from image scoring, whereas there was no such trend when only partial information is provided. Our results are thus in the line with these two hints from Milinski et al. (2001) in a different way of speaking.

In sum, our results are very robust and consistent with the past studies, some general discussions and conclusions will be made in the last chapter.

# Conclusion

As we have explained in the first chapter, plenty of past literatures have theoretically argued the importance on the evolution of indirect reciprocity since the real society involves much more complicated scenarios rather than two same actors interact all the time. As in the real life, there possibly has very limited information about your partner if you only will meet him/her once maximally in your life. But perhaps direct or indirect information can be available at the certain point/level, which can build up the first- and second-order information about each individual. The main disagreement nowadays occurs at the point of determination on whether treat the ones give to 'bad' as 'bad' in the indirect reciprocity. Ohtsuki and Iwasa (2004) argued that the definition of goodness should include the individuals who give to 'bad', whereas Takahashi and Mashima (2006) argued that such piece should not be included in the definition of goodness. Despite of that, all the relevant experimental studies showed that the both first- and second-order information should be taken into account when facing a cooperation dilemma. The remaining puzzle in our mind is whether people do actually make information storage and recall when it is necessary for the indirect reciprocity.

The helping experiment we conducted is modified from Bolton et al. (2005); it shows that by providing additional second-order information, the probability of cooperation

does increase comparing with only the first-order information is present. And information condition, session, or round has very significant effect on the probability of cooperation. The effect of condition and round is moderated by session because it is quite obvious that at the end of each session the rate of cooperation decreases, and both the FI and PI condition have a decreasing trend but not in the NI condition. Those all implies that people usually less cooperative when the end of the session approaches. Such behavior can be quite realistic, because person 1 would only be paired with person 2, for instance, for one session. As the ending rounds coming up, there is little reason for them to cooperative if they think such reputation would not be taken into account. But those outcomes are only based on an overall looking at the rate of cooperation.

After that, we distinguish four different cases that might occur – CG, CB, DG, and DB in order to gain a better picture on the action and reputation mechanism. When collapsing the sessions, it is more likely that people are adopting image scoring strategy in the full information condition; strict discriminator strategy in the partial information condition; and All-D strategy in the no information condition. It implies that people do evaluate giving to 'bad' differently in different information condition. And under the FI condition, people intend to take both first- and second-order information into account when they are making choices on cooperation, people seems to try to remember as much as information under the PI condition. It shows that

people are at a better information storage and recall situation when there is only first-order information presenting. The rough significance for the usage of second-order information further indicates that even though the second-order information is available, people do still tend to remember as much as they could on the first-order information. While the choices made in the NI condition, is less likely can be reasoning, however, we did observed certain level of cooperation in the NI condition, which implies that not all the cooperation are attributable from the image scoring strategy. The probability of cooperation in CG and CB case are more likely to be higher than the rate in DG and DB case in general. The trends of all four cases are more likely to be declining.

Nevertheless, we find the evidence that people intend to storage/recall both first- and second-order information due to the fact that by offering additional second-order information the rate of cooperation increases, but with a rather marginal level of significance for using of the second-order information. As a result, it is not conclusive that if people do actually make information storage and recall when it is needed, simply because such increase in the rate of cooperation could be resulting from the illusion effect that people just tend to cooperate more when more information is available. That is the reason we conducted further step along with the post questionnaire. In the post questionnaire, candidates were asked four general question (for all three information conditions) to see the seriousness of earning money, whether maximization of self-earnings, the importance of the whole population's wealth, and

competition between self-earning and others' earnings have some impacts on the decision of cooperation ; then they were exposed to the question concerning the information storage, caring about past behavior and past treatment by the others in the first two information conditions; as well the level of information recall only during the first three sessions. Results show that the behavior and intuition claimed in the post questionnaire is often corresponding to how people actually did do during the experiment. The group effect is not that much observed, only at the question about the degree of caring of the past treatment from the others under the FI condition.

People, who are seriously aiming at earning money, are more likely to pay attention on the second-order information, and try to make information storage on both first- and second-order information. Even though the interaction effect that including both action and reputation did not show significance, we suspect the possible cause can be storing both first- and second-order information accurately and retrieve from memory in the later on rounds perhaps require too much memory capacity.

For the wealth maximization, quite surprisingly see that maximization of the whole group's wealth is the only factor has significance when interacts with reputation. It indicates that people are actually trying acting as socially responsible. That means that players have more incentives on increase the whole group's welfare/payoff rather than own interest on earnings or the gap between self- and others' earnings.

For the information storage and recall, we did found evidence about the storage, but not that much on recalling. And more storage and recalling are significant when there is only the first-order information or only the first-order information is been put into the memory even though it is in the full information condition. On the other hand, the ones who care more about people's past are tend to cooperate in a more justifiable way. The similar suspicion mentioned earlier about the too much memory capacity can be the reason. The thing is individuals are actually trying to put the information into their own memory. And out suspicion reasoning is quite realistic, because we observed that when there is only first-order information available, the information about the givers' action has been stored and recalled more. Such results can also be the case of the Dutch sample, which leads to the possible further analysis for a more international and diversified sample.

As we had already mentioned, the strategy people adopting is more in the line with image strategy, or SDICS strategy, but not the standing strategy. It might be due to the fact that making information storage into memory and recall when the later on rounds start is requiring lots of demanding on the memory capacity. But people do have different characteristic on this issue, which could cause less possibility for them to adopt the standing strategy or more justifiable strategies. As a result, the strategies adopting for the individuals so far shown in our experiment is sharing the same common feature: (1) treating the ones giving to 'good' as 'good'; (2) treating the ones not giving to 'good' as 'bad'; (3) treating the ones not giving to 'bad' as 'bad'. In this

sense, we are having a very robust and consistent outcome with the previous theoretical and experimental literature. The first two features are also very making common social sense.

However, due to the certain constraint, we did not run all the variables we do have , for some test that some previous studies have conducted, for instance, whether the earnings and total time taken have an effect on people to make decisions on the cooperation. Those will be addressed in the further research agenda.

# References

Andreoni, J., and J. Miller (2002), Giving According to GARP: An Experimental Test of The consistency of Preferences for Altruism. *Econometrica*, Vol. 70, 737-753

Axelrod, R., and W. D. Hamilton (1981), The Evolution of Cooperation. *Science*, Vol. 211, 1390-1396

Bolton, G. E., E. Katok, and A. Ockenfels (2005), Cooperation among Strangers with Limited Information about Reputation. *Journal of Public Economics*, Vol. 89, 1457-1468

Engelmann, D., and U. Fischbacher (2002), Indirect Reciprocity and Strategic Reputation Building in An Experimental Helping Game. *Working Paper Series_Institute for Empirical Research in Economics University of Zurich*, ISSN 1424-0459, Working Paper No. 132

Fischbacher, U. (2007), z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics*, Vol. 10, 171-178.

Greiner, B., and M. V. Levati (2005), Indirect Reciprocity in Cyclical Networks: An Experimental Study. *Journal of Economic Psychology*, Vol. 26, 711-731

Mashima, R., and N. Takahashi (2006), The Importance of Subjectivity in Perceptual Errors on The Emergence Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 243, 418-436

– (2007), *The Emergence of Generalized Exchange by Indirect Reciprocity*. Springer. (Chapter 10 of the New Issues and Paradigms in Research on Social

Dilemmas by A. Biel et al., 159-176)

Milinski, M., D. Semmann, T. C. M. Bakker, and H.-J. Krambeck (2001), Cooperation Through Indirect Reciprocity: Image Scoring or Standing Strategy? *Proceedings of the Royal Society of London Series B: Biological Sciences*, Vol. 268, 2495-2501

— (2002), Donors to Charity Gain in Both Indirect Reciprocity and Political Reputation. *Proceedings of the Royal Society of London Series B: Biological Sciences*, Vol. 269, 881-883

Nowak, M. A., and K. Sigmund (1998), The Dynamics of Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 194, 561-574

— (2005), Evolution of Indirect Reciprocity. *Reviews_Nature Publishing Group*, Vol. 437, 1291-1298

Ohtsuki, H. (2004), Reactive Strategies in Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 227, 299-314

Ohtsuki, H., and Y. Iwasa (2004), How Should We Define Goodness? – Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, Vol. 231, 107-120

— (2006), The Leading Eight: Social Norms That Can Maintain Cooperation by Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 239, 435-444

Panchanathan, K., and R. Boyd (2003), A Tale of Two Defectors: The Importance of Standing for Evolution of Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 224, 115-126

Seinen, I., and A. Schram (2001), Social Status and Group Norms: Indirect Reciprocity in A Helping Experiment. *Working Paper from University of Amsterdam*, 1-23

Suzuki, S., and E. Akiyama (2008), Chaos, Oscillation And The Evolution of Indirect Reciprocity in N-Person Games. *Journal of Theoretical Biology*, Vol. 252, 686-693

– (2008), Evolutionary Stability of First-Order-Information Indirect Reciprocity in Sizable Groups. *Theoretical Population Biology*, Vol. 73, 426-436

Takezawa, M. (2009), Indirect Reciprocity with The Information Recalled from The Memory: An Experimental Study. (Presentation PowerPoint), 1-46

Tanimoto, J. (2007), A Study of Indirect Reciprocity Involving A Reputation System or A Simple Tag System in A One-Shot, Multi-Player Game. *Biology Systems*, Vol. 90, 856-869

Uchida, S., and K. Sigmund (2010), The Competition of Assessment Rules for Indirect Reciprocity. *Journal of Theoretical Biology*, Vol. 263, 13-19

Wedekind, C., and V. A. Braithwaite (2002), The Long-Term Benefits of Human Generosity in Indirect Reciprocity. *Current Biology*, Vol. 12, 1012-1015

Wedekind, C., and M. Milinski (2000), Cooperation Through Image Scoring in Humans. *Science*, Vol. 288, 850-852

## Table 1 Four Genes Determining Strategies

| | | Second-order information *Reputation* | |
|---|---|---|---|
| | | **'Good'** | **'Bad'** |
| First-order Info. *Action* | **Cooperate** | Gene1: good/bad | Gene2: good/bad |
| | **Defect** | Gene3: good/bad | Gene4: good/bad |

Note: First-order information refers to current recipient's behavior towards the previous recipient; while second-order information refers to current recipient's previous recipient's score.

In total, 16 strategies can be accounted. For example, ALLC is represented as GGGG, ALLD is represented as BBBB, Image scoring strategy as GGBB, standing strategy as GGBG, extra standing strategy as GBBG, SDISC as GBBB, and ES as GBBG.

Source: Takahashi and Mashima, 2006: 420

## Table 2 Probability of Cooperation for Different Situations

### (a) Full Information Condition

| | Reputation | |
|---|---|---|
| | **Good** | **Bad** |
| **Cooperate** | .40<br>(0.20, 85) | .37<br>(0.26, 85) |
| **Defect** | .30<br>(0.22, 85) | .26<br>(0.22, 85) |

### (b) Partial Information Condition

| | Reputation | |
|---|---|---|
| | **Good** | **Bad** |
| **Cooperate** | .34<br>(0.21, 84) | .28<br>(0.26, 84) |
| **Defect** | .21<br>(0.20, 84) | .21<br>(0.20, 84) |

**(c) No Information Condition**

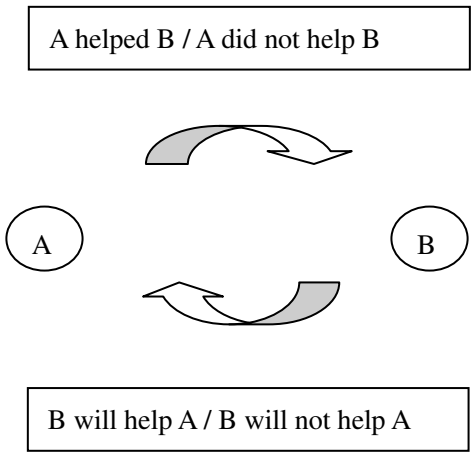
| | Reputation | |
|---|---|---|
| | **Good** | **Bad** |
| **Cooperate** | .19 | .19 |
| | (0.22, 77) | (0.25, 77) |
| **Defect** | .17 | .18 |
| | (0.18, 77) | (0.21, 77) |

Note: Numbers in parentheses refer to standard deviation and number of cases, respectively.

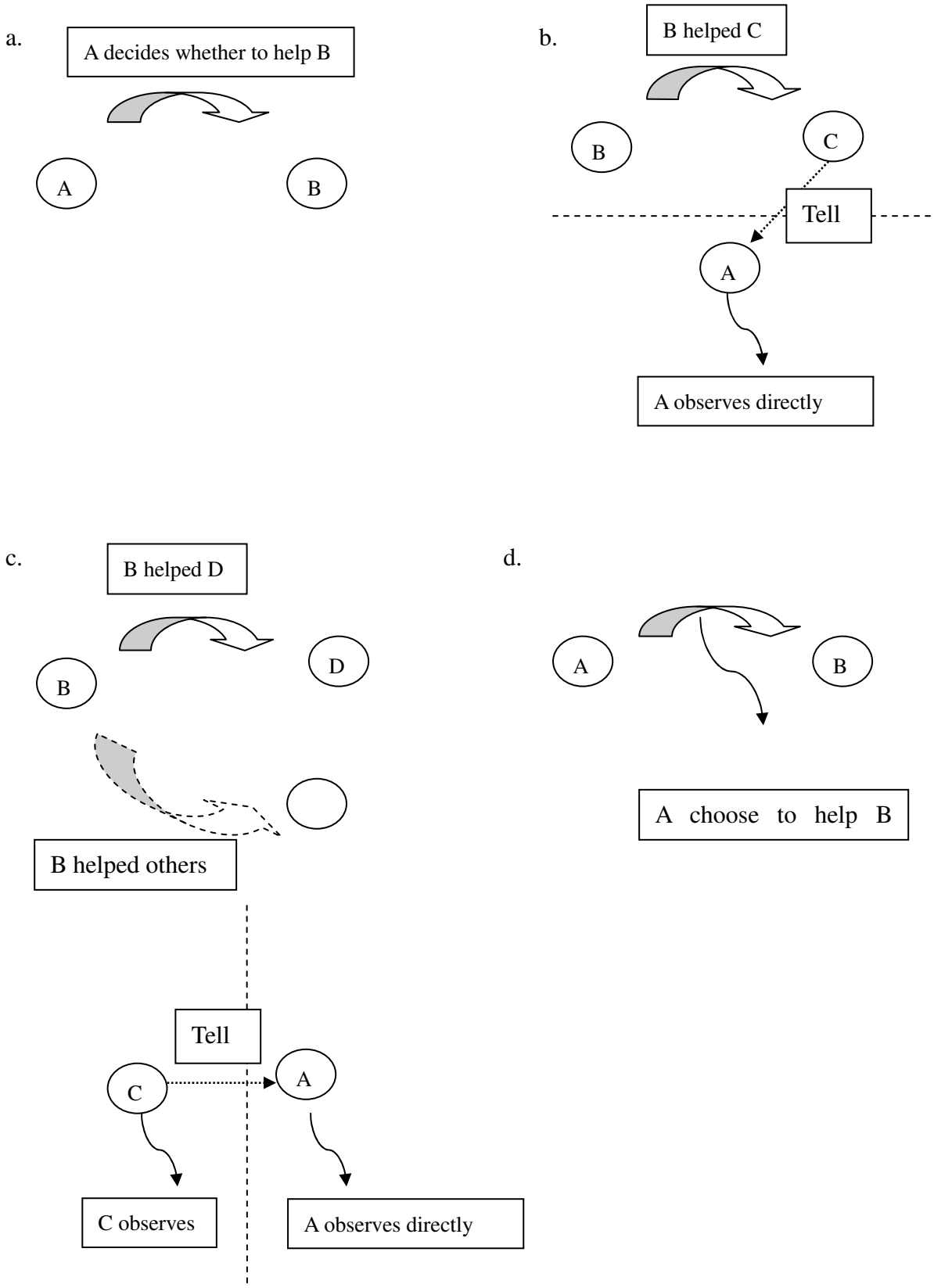
## Table 3 The Impact of Receiver's Histories on the Probability of Cooperation

| Information condition | | FI | PI | NI |
|---|---|---|---|---|
| | **DB** | .28 | .26 | .26 |
| *Receiver's last move* | **DG** | .25 | .21 | .25 |
| *(second-order information)* | **CB** | .35 | .31 | .22 |
| | **CG** | .36 | .32 | .26 |

## Figure 1 Direct Reciprocity

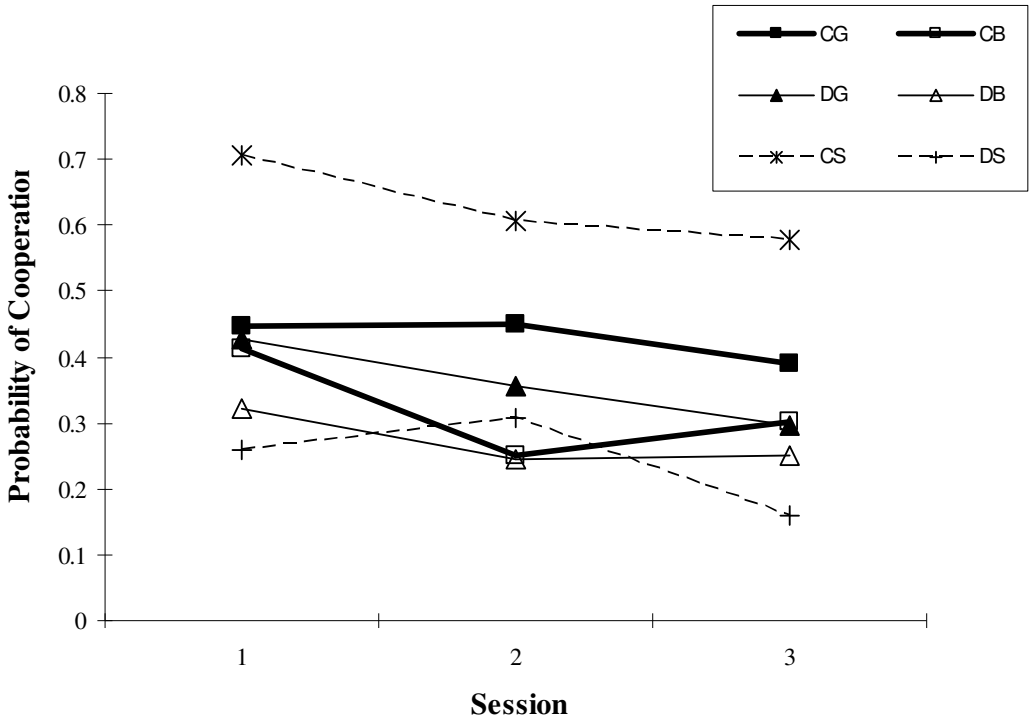


## Figure 2 Indirect Reciprocity

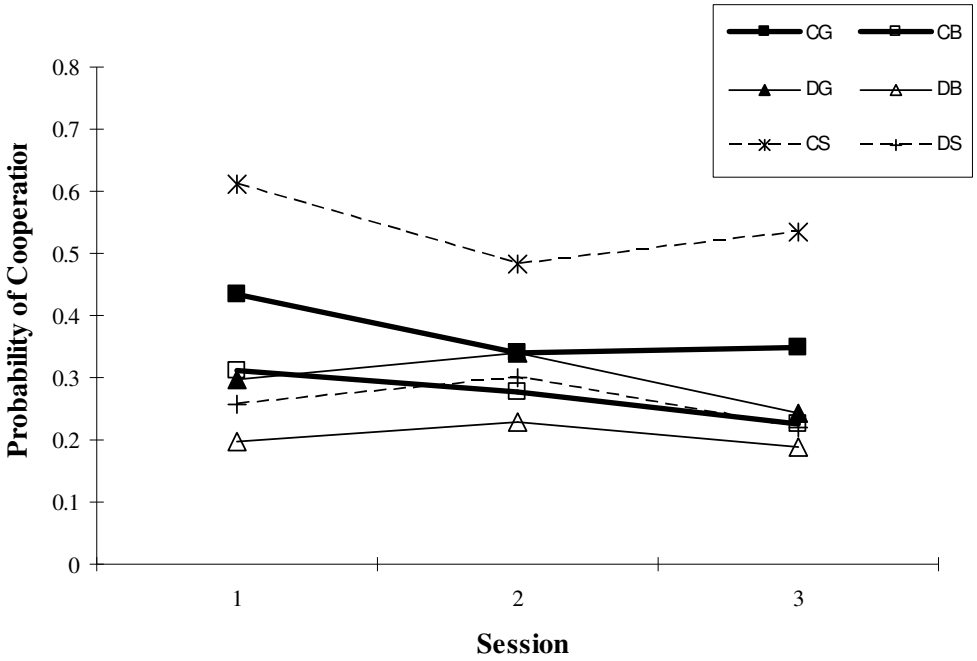


# Figure 3 Average Probability of Cooperation for All 6 Cases

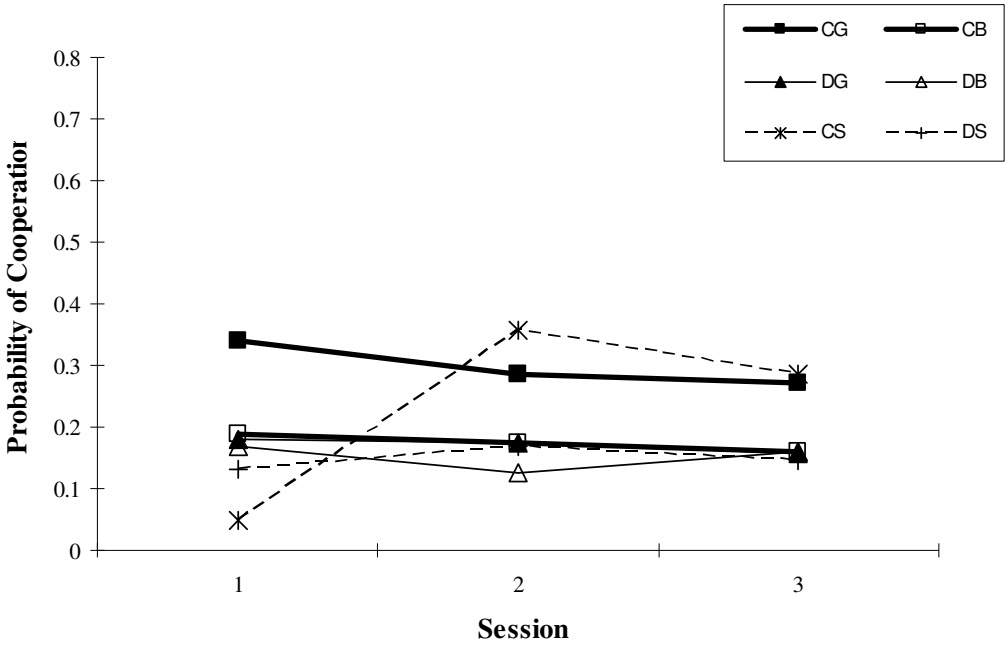
## (a) Full Information Condition



**(b) Partial Information Condition**

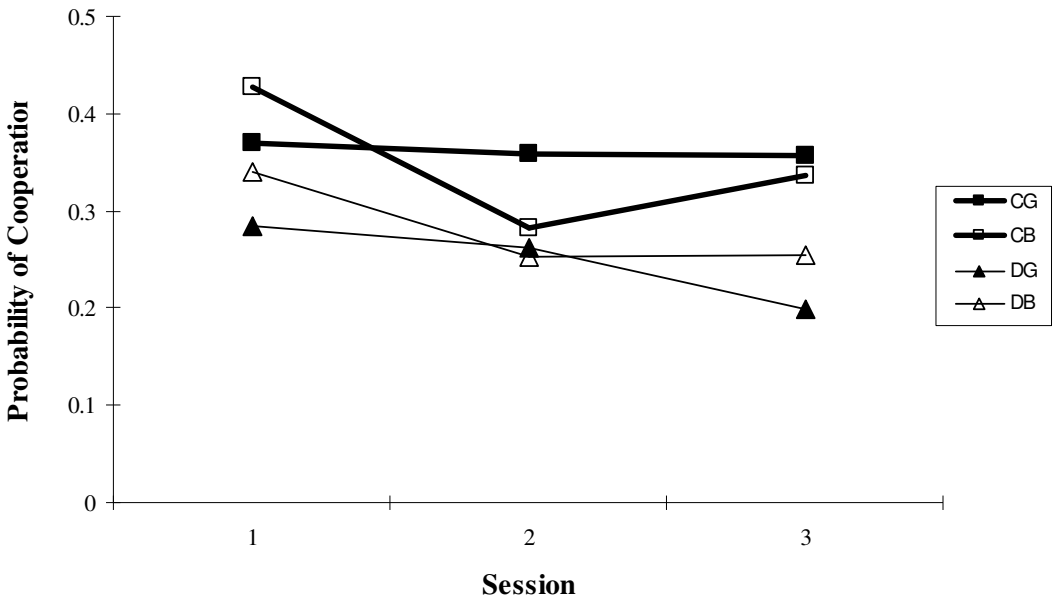


**(c) No Information Condition**

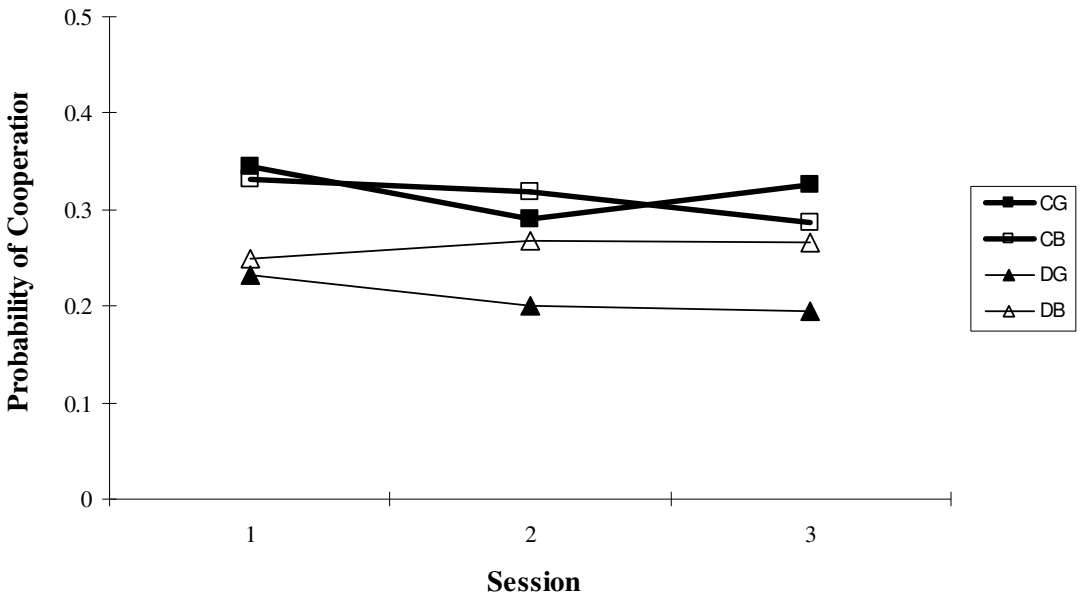


# Figure 4 Average Probability of Cooperation per Condition

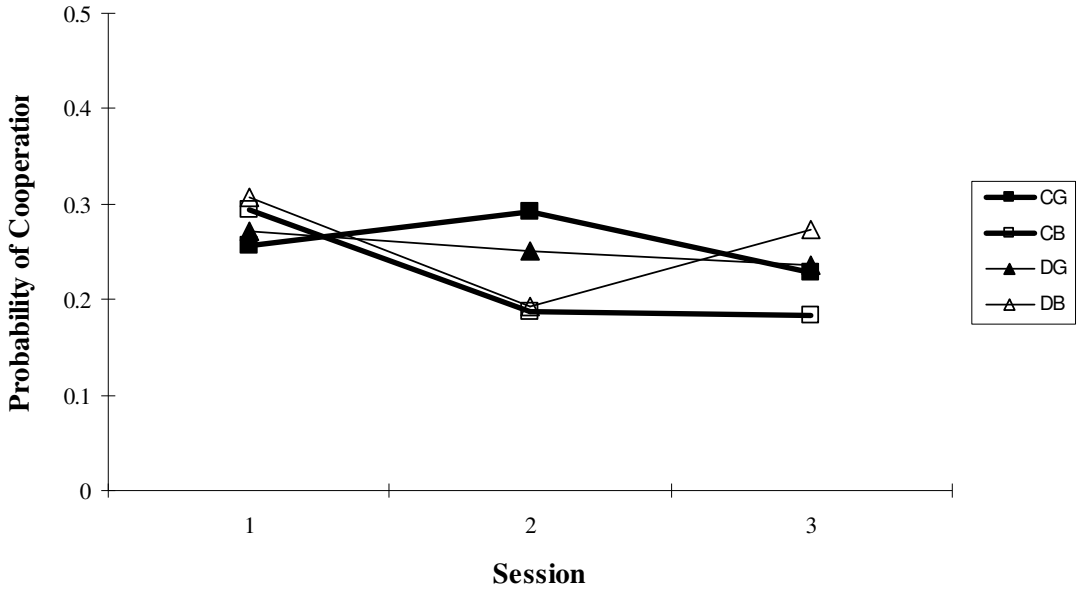
## (a) Full Information Condition



**(b) Partial Information Condition**

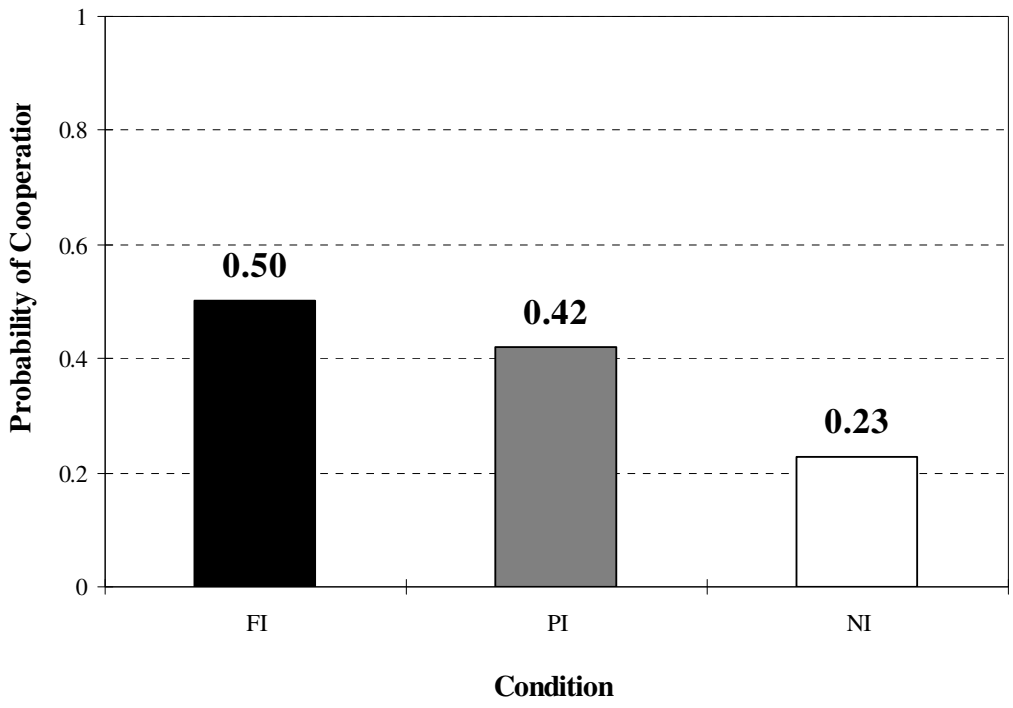


**(c) No Information Condition**

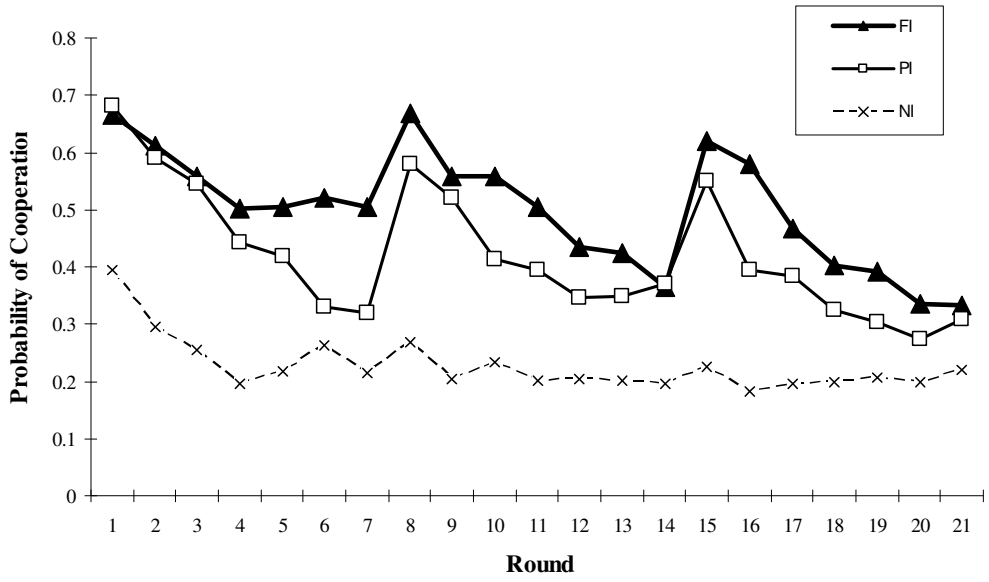



Note: The difference between Fig. 3 and 4 is that Fig. 4 is based on the results from the GLM.

## Figure 5 Probability of Cooperation for Each Condition



# Figure 6 Trends of Probability of Cooperation



# Acknowledge

This piece of master thesis is the output that involved many people's efforts – my supervisor, families, friends, professors, colleagues from FSW to participate in our experiment, as well colleagues from GCC. I indeed feel much appreciated about their sharing of knowledge and comments with me. That is the main drive of mine to go further every little step until the end.

It actually did take quite a long period of time for my supervisor and me to decide to draft such paper in this certain topic. Initially, we were planning to write on the topic of regret and inaction. Somehow, we turned out to put the interest onto the indirect reciprocity in the end. I am actually glad that I did finish my master thesis on current topic, which showed me a lot of insights about real life behavior and social networks.

Frankly speaking, I am a student from the Faculty of Economics and Business (FEB) in Tilburg University, however, I am really grateful to have such opportunity to get to know how research has been conducted in the Faculty of Social and Behavioral Sciences (FSW).

Most importantly, I would love to thank my supervisor professor M. Takezawa, who has spent countless time on reading and correcting my drafts. In particular, I did learn directly how to write up a proper research paper due to the fact that he always makes

the checking together with me, shows me how the right way should be. Even though this is the very first time for me to do research on psychology and behavior economics, I do think I made the right choice to take such a challenge. Dr. M. Takezawa opened a new door for me to access the new area of expertise, I do appreciate it a lot. His expert knowledge and feedback are the crucial part of completing this piece. He also encourages me to think actively on the topic, which let me enjoy the fun part of doing this master thesis.

As well, I would love to give special thanks to Maartje Elshout, Willem-Jan, and Job van Wolferen from the ESP lab, who offered me lots of support when I was conducting the experiments. Simply because this is entirely first access for me, their help is essentially important for me to do the right steps and finish the whole progress.

Some special thanks to my families, other friends, colleagues and professors, they have been there for me during the whole process. Especially my uncle and aunty, who helped me for checking the entire paper.

Without one of them, I will not be able to achieve the final destination of this paper. I am truly thankful to all the supports I have received.

As the final point, I also owe some thanks to Professor Charles Noussair, who as the second reader for my master thesis has put into lots of hours reading my paper.

Thanks a lot for all of you very much!

# Appendix

## I. Instructions for the Experiment

(1) How to Make Decisions

Three assigned givers and three assigned receivers in every round.

Following screen will be showing at the beginning of each round. Take an example, you are player 4.

| Please make a choice for each player. | | |
|---|---|---|
| Player 1 | ○ give the money | ○ don't give the money |
| Player 2 | ○ give the money | ○ don't give the money |
| Player 3 | ○ give the money | ○ don't give the money |
| Player 5 | ○ give the money | ○ don't give the money |
| Player 6 | ○ give the money | ○ don't give the money |
| Player 7 | ○ give the money | ○ don't give the money |

At current moment, you have no idea who will be assigned as a giver, or who will be assigned as a receiver. You need to ask yourself whether you would like to help Player 1 for instance if you are in the position of the assigned giver.

You need to repeat this decision making process for all the 6 other people.

Basic settings are identical for all three conditions. The differences are only on the information/feedback controlling part. See more details in the following.

(2) Result screens for FI

Player 3 gives the money to Player 5

| |
|---|
| Player 4 doesn't give the money to Player1<br><br>Player 6 gives the money to Player 2 |

This information is displayed for a while and automatically disappears.

Next, a screen with the earnings of yours for so far in the current session will be displayed.

After that, a new round will automatically start.

(3) Result screen for PI

| |
|---|
| Player 3 gives the money<br><br>Player 4 doesn't give the money<br><br>Player 6 gives the money |

This information is displayed for a while and automatically disappears
.
Next, a screen with the earnings of yours for so far in the current session will be displayed.

After that, a new round will automatically start.

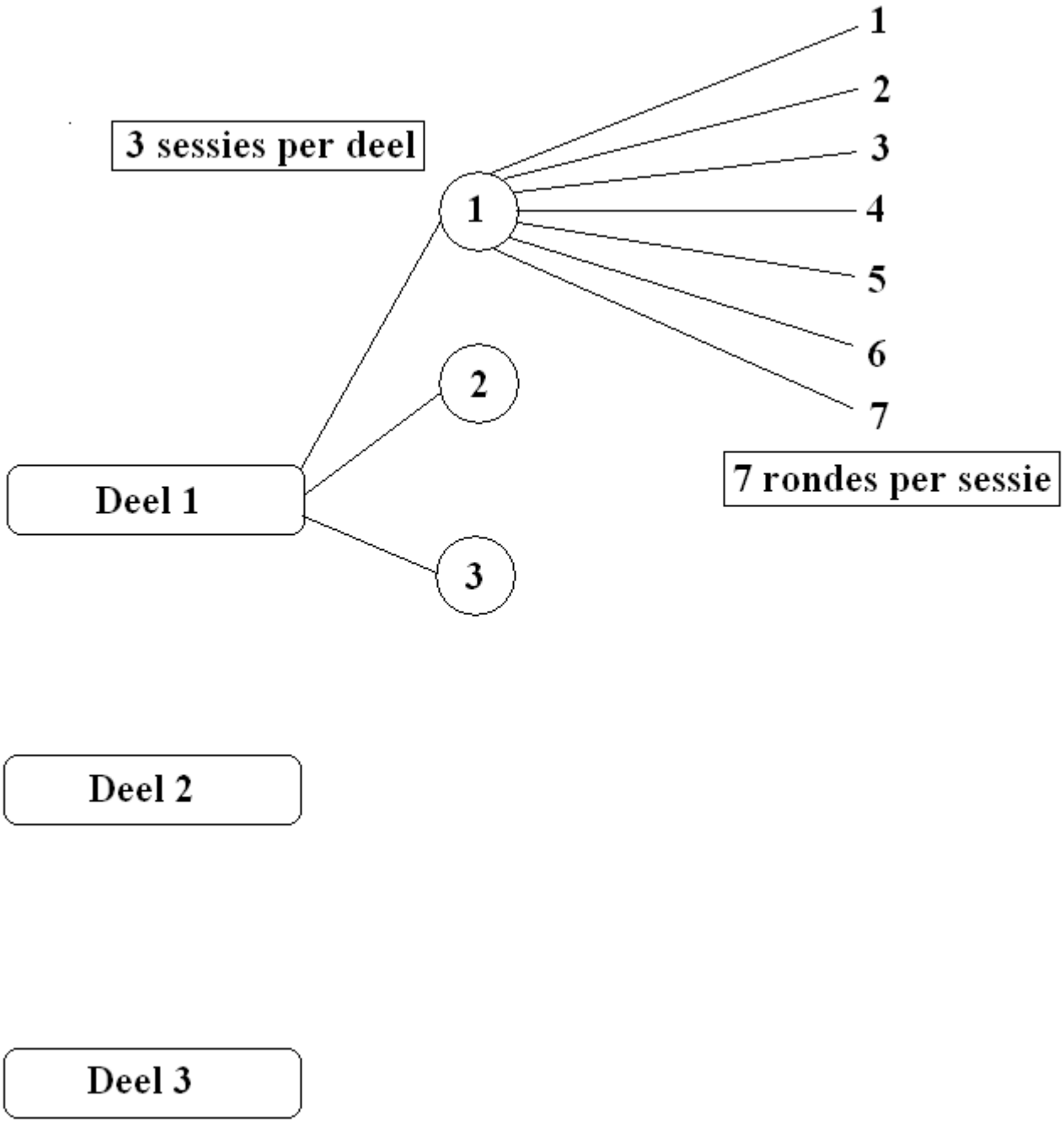
(4) Results screen for NI

We do NOT show you neither assigned giver ID nor assigned receiver ID.

We only show you:

- Whether you are selected as an assigned giver or receiver
- Whether you have been helped by the others if you are assigned as a receiver
- Total amount of earnings in the current session

(5) Overall structure



## II. Post Questionnaires

What is your gender? Male / Female

What is your age? [ ]

| | | | | | |
|---|---|---|---|---|---|
| How serious were you interested in earning the money during the experiment? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Did you try to maximize only your earning? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Did you try to maximize the sum earning of all the people? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Did you try to maximize the difference between your earning and the average earnings of the others? | not at all 1 | 2 | 3 | 4 | totally 5 |

***Following, we want to ask some questions about what you had thought during the first 3 sessions.***

| | | | | | |
|---|---|---|---|---|---|
| Q1. Did you try to remember as much as you can when seeing the results of decisions made by three GIVERs? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q2. When making decisions, did you care the past behavior of the others? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q3. When making decisions, did you care how the other people treated you in the past? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q4. When you saw the information GIVER gave (did not give) the money to RECEIVER, did you consider whether or not RECEIVER had (not) given the money to the others when they were selected as GIVER in the past? | not at all 1 | 2 | 3 | 4 | totally 5 |

***We want to ask some questions about what you had thought during <u>the second 3 sessions.</u>***

| | | | | | |
|---|---|---|---|---|---|
| Q1. Did you try to remember as much as you can when seeing the results of decisions made by three GIVERs? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q2. When making decisions, did you care the past behavior of the others? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q3. When making decisions, did you care how the other people treated you in the past? | not at all 1 | 2 | 3 | 4 | totally 5 |

***Now, we want to ask some questions about what you had thought during <u>the third 3 sessions.</u>***

| | | | | | |
|---|---|---|---|---|---|
| Q1. When making decision, did you care how the other people treated you in the past? | not at all 1 | 2 | 3 | 4 | totally 5 |
| Q2. When you see the result screen (showing how much you earn in each round), did you consider to change your behavior only based on the earning results? | not at all 1 | 2 | 3 | 4 | totally 5 |