Cooperation in Prisoner's Dilemma Game: Influence of Social Relations

Maurice Grinberg (mgrinberg@nbu.bg)
Evgenia Hristova (ehristova@cogs.nbu.bg)
Milena Borisova (borisova_milena@abv.bg)

Department of Cognitive Science and Psychology,
New Bulgarian University, 21 Montevideo Street, Sofia 1618, Bulgaria

Abstract

The paper explores the influence of the type of relations among players on cooperation in the Prisoner's dilemma game. The relations between players are operationalized according to Fiske's relational models theory (Fiske, 1991): communal sharing, authority ranking, equality matching, and market pricing. This is achieved by using various ways of distributing the total payoff gained by a dyad of players in a series of Prisoner's dilemma games: each player receives the total payoff (*unity*), one of the players receives more than the other (*hierarchy*), each player receives half of the total payoff (*equality*), each player receives a portion of the total payoff proportional to his/hers individual payoffs (*proportionality*). For these four conditions, the cooperation rates, the mutual cooperation, the mutual defection, and the payoffs gained are analyzed and compared for a series of forty games. The results show that in the *proportionality* condition there is less cooperation, less mutual cooperation, more mutual defection and less total payoff than in the other three conditions.

Keywords: Prisoner's Dilemma, decision-making, cooperation, social interaction, relational models

Introduction

Prisoner's Dilemma Game

The Prisoner's dilemma (PD) game is one of the most extensively studied social dilemmas. PD is a two-person game. The interest in studying PD game arises from the idea that many social situations and problems such as overpopulation, pollution, energy savings, participation in a battle, etc. have such a dilemma structure (Dawes, 1980). The payoff table for this game is presented in Figure 1. In the PD game the players simultaneously choose their moves – C (cooperate) or D (defect), without knowing their opponent's choice.

In order to be a Prisoner's dilemma game, the payoffs (see Figure 1) should satisfy the inequalities T > R > P > S and 2R > T+S. Because of this game structure a dilemma appears – there is no obvious best move. On one hand, the D choice is dominant for both players – each player gets larger payoff by choosing D (defection) than by choosing C (cooperation) no matter what the other player chooses. On the other hand, the payoff for mutual defection (P) is lower than the payoff if both players choose their dominated C strategies (R for each player).

As PD game is used as a model for describing social dilemmas and studying the phenomenon of cooperation, there is a great interest in the conditions that could promote or diminish cooperation.

Player I	*Player II* C	*Player II* D
C	R, *R*	S, *T*
D	T, *S*	P, *P*

Player I	*Player II* C	*Player II* D
C	3, *3*	1, *4*
D	4, *1*	2, *2*

Figure 1: Payoff tables for the PD game – with standard notation for the payoffs and an example. In each cell the comma separated payoffs are the Player I's and *Player II*'s payoffs, respectively.

In game theory several assumptions about the game and the players are made. The agents are assumed to be perfectly rational and to have perfect information about the game. Under these conditions, they are supposed to try to maximize their payoffs in a completely selfish manner (Colman, 2003). From this point of view the dominant strategy in the game is defection (in one-shot or in repeated PD games with a fixed and known number). This prediction is in contrast with the behavior of the players observed in laboratory settings or in real life situations.

In human societies, people cooperate all the time and often cooperation is seen as one of the foundations of human civilization (see e.g. Gärdenfors, 2003). Sally (1995) provides a meta-review of the experiments involving PD games published between 1958 and 1995 and shows that in its iterated version (the game is played many times), cooperation choices are made in 20-50 % of the games (mean 47.4 %) and even in one-shot games many players cooperate although much less than in the iterated version.

Several studies have shown how cooperation can emerge from expected utility or anticipatory reinforcement models without any specific relations between the players (Grinberg, Hristova, & Lalev, 2010; and the references there in).

However, it is clear that the deeper understanding of how people make decisions while playing PD games should account for the role of the social relations involved in the interactions. Moreover, as the PD game is central in the modeling of social interactions it can be used to explore the existence and limits of the relational social types as posited by relational social models (see e.g. Haslam, 2004). Exploring the potential of games like the PD game as modeling relational types is one of the goals of this paper.

Social Interactions and Cooperation

How decision-making in PD games is influenced by social interactions has been explored in many studies that try to account for the contradiction between the normative predictions and the experimental results in PD games.

Several studies have established the influence of social interaction on cooperation. For instance, Durkin, Frost, Aronov & Breslow (1967) found that cooperative moves double when participants have visual contact with each other compared to the condition where they don't have such contact. Sally (2001) investigated the behavioural changes in participants who know each other or are psychologically or socially close and discussed the importance of such closeness in game strategy building. According to this account, participants play differently depending on how they perceive their opponent – as a friend or a stranger. In both cases, according to Sally (2001), the social interaction is essential and the social dilemmas like PD need to be investigated from the perspective of a general relational theory.

There are several theories that account for the cooperative behavior in PD games in terms of socially established values and stress the importance of social interaction and relationships as tools for achieving cooperation. Among them are theories that explain cooperation by altruism, reciprocity or reputation building.

Reputation building theory (Kreps et al., 1982; Andreoni & Miller, 1993) is one of the main theories aimed to explain cooperation in iterated PD game. This theory assumes that players are self-interested (not altruists), but the repetition in iterated PD games creates incentives to cooperate. According to this model, the player is building himself a reputation of a cooperative player and expects that the other player will also cooperate.

Reciprocity, according to many researchers, is a widespread norm and is the basis of many relationships and societies (Trivers, 1972). People reciprocate cooperation with cooperation. One of the most studied strategies that are based on reciprocity is the tit-for-tat strategy. A player using this strategy cooperates initially, and then plays the same as his/her opponent did in the previous game. It has been demonstrated in computer tournaments that in the long run the tit-for-tat strategy results in higher payoffs compared to other strategies (Axelrod & Hamilton, 1984; Komorita, Hilty, & Parks, 1991).

Another influential theory about cooperation in PD game is based on the concept of *altruism*. In contrast to reputation building theory, this theory assumes that some players are not strictly self-interested and view more benefit in cooperation than the actual payoffs they receive (Cooper et al., 1996). From an altruistic perspective, cooperation can yield higher payoffs than defection.

Although these social theories of cooperation have been proposed to explain cooperative behavior unexpected by normative game theory, it is interesting to consider more general social theories that are more closely related to the game theoretic analysis of social relations. In our opinion such a theory is the relational models theory proposed by Alan Fiske (Fiske, 1991).

Relational Models Theory

Relational models theory (Fiske, 1992; Fiske & Haslam, 1996; McGraw, Tetlock, & Kristel, 2003; Rai & Fiske, 2011) states that there are four basic schemas that are used to build, organize and maintain relationships and interactions among individuals in a society. These models are supposed to be universal and all relations could be described by these models or by combination of them. The four types of relations generate four modes for every aspect of the interactions between people – resource allocation, moral judgments, decision-making, etc.

These four relation models are the following (Fiske, 1992):

- **Communal Sharing** – relations in an undifferentiated group of people with equivalent status. Everyone in a community - which could consist of two members or could be very large – has some rights and some duties. The focus is on commonalities and not on distinctions;
- **Authority Ranking** – implies an ordinal ranking in society and this ranking scheme determines one's relative status. For instance, military hierarchy can be considered a prototype of such relations;
- **Equality Matching** – relations are based on a model of one-to-one correspondence as in turn-taking, tit-for-tat strategies, etc. The social prototype would be friendship networks, in which reciprocity is a norm which rules the distribution of wealth;
- **Market Pricing** – based on a model of proportionality in social relations in which people reduce their interaction to some ratios of utility measures. Examples of relations of this type are the ones governed by prices, rational calculations, expected utilities, etc.

Payoff Distribution in PD and Fiske's Relational Model Types

Fiske's relational model theory (Fiske, 1992) claims that different relational models influence and are manifested in a lot of domains and activities, e.g. reciprocal exchange, distribution, contribution, work, significance of time, social influence, constitution of groups, motivation, moral judgments, etc.

Here, we focus on the type of distribution of group resources using one and the same game, namely the PD game. We focus on social interaction related to social exchange as instantiated by contribution and distribution of a common resource. We share the opinion that situations involving exchange are the most appropriate to study the four types of relational models in isolation (Haslam, 2004).

In the classical PD game experiments each player is rewarded according to his/hers personal payoffs. However, Fiske's relational model theory states that in real-life situations the distribution of payoffs and resources depends on the type of the relational model behind the social interaction. There are **four types of distributions**, corresponding to the four relational types described above (Fiske, 1992, Table 1, p. 694):

- **communal sharing** – 'corporate use of resources regarded as common, everything belongs to all together';
- **authority ranking** – 'the higher the person's rank, the more he or she gets';
- **equality matching** – 'to each the same, everyone gets identical shares';
- **market pricing** – 'to each in due proportion'.

Goals of the Study

The main goal of the present study is to make a first step in the mapping of Fiske's relational models theory to games from game theory. More specifically, we want to study how the four relational types, implemented as distinct payoff distributional models, influence cooperation in PD games. As relational models are complex and encompass various domains, in the present study the focus is on the different **distribution schemas** within the same type of games (the PD game).

We aim to explore what is the influence of the type of relation among players on a set of game outcomes that characterize the playing of a PD game – cooperation, mutual cooperation, and mutual defection. It is also important to check the influence of the distribution model on the overall payoffs that are received – e.g. what type of model is more beneficial in terms of total payoff earned in interactions with the strategic structure of the PD game.

Cooperation is expected to be the highest if the payoff distribution is in accordance with the communal sharing model. Cooperation is expected to be lowest if the distribution follows the rules of market pricing model, e.g. when everyone is rewarded depending on his/her personal contribution – in this scenario we expect more individualistic orientation of the players.

Method

Stimuli

A sequence of 40 Prisoner's dilemma games is used in the experiment. All of the games used had the payoff matrix given in Figure 2. At the beginning of the series there were 5 training games (results from these games are not included in the analysis) thus the total sequence comprised 45 games.

Player I		*Player II* C	*Player II* D
	C	40, *40*	10, *50*
	D	50, *10*	15, *15*

Figure 2: Payoff table for the PD game used in the experiment.

Experimental Conditions

The **distribution of the total payoff** is varied in accordance with the four relational models described above in a between-subjects design. There are four experimental conditions that differ in the way that the total payoff of a pair is divided between the players in that pair:

- **Unity condition** – each player receives the total payoff earned by the pair (*communal sharing* relational model);
- **Hierarchy condition** – one of the players receives more than the other – 2/3 vs. 1/3 of the total payoff of the pair (*authority ranking* relational model);
- **Equality condition** – each player receives equal portion of the total payoff (*equality matching* relational model);
- **Proportionality condition** – each player receives a share of the total payoff proportional to his/hers individual payoffs (*market pricing* relational model).

Procedure

Subjects were tested in pairs. After receiving the appropriate instructions for the experimental condition they were in, each dyad played 5 training games, followed by 40 games that were analyzed. The experimenters secured that the participants will not have visual, verbal and any kind of other contact before and during the experiment. Therefore, no player knew who the other player was before the end of the experiment.

Instructions for the experiment explained in details the rules of the game and included several test questions to make sure that the participants understood correctly the rules. There were four instructions that varied only in the explanation for the total monetary payoff distribution. They are quoted below because they define the relational models in the four conditions:

- *Unity condition* – 'Each of you will receive the amount of money you have earned together';
- *Hierarchy condition* – 'You will get 2/3 of the total amount of money of the pair, and the other player will get 1/3 of it (for one of the players). You will get 1/3 of

the total amount of money of the pair, and the other player will get 2/3 of it (for the second player)';

- *Equality condition* – 'The total amount of money of the pair will be split equally between you and the other player';
- *Proportionality condition* – 'The total amount of money of the pair will be split between you and the other player in accordance with the number of points each of you has earned'.

The game was presented in a formal and a neutral formulation. On the interface, the cooperation move was labeled '1' and the defection move was labeled '2'. However, further in the paper, for convenience, we will continue to use *cooperation* instead of move '1' and *defection* instead of move '2'. Matlab 7.6.0 (R2008a) was used for presenting the game and recording the choiches of the players.

After each game the subjects got feedback about their own and the other player's choice and payoffs in the current game. They could also constantly monitor their own total payoff; the total payoff of the other player; the total payoff for the pair, and the monetary equivalent of the total payoff of the pair (that is to be distributed among them).

Participants were instructed to try to *maximize* the amount of money they will get. Subjects were paid real money accordingly to the final payoff in the game. Players in the *unity* condition received the same amount of money for 1000 points as participants in the other 3 conditions received for 500 points. Thus we tried to equate the absolute magnitude of the monetary payoff that the participants could receive during the experiment.

Each session lasted about 20 minutes.

Participants

80 participants (47 female, 33 male) took part in the experiment. All of them were university students, mean age 23.3 years.

They were tested in 40 pairs – 10 pairs in each experimental condition. Participants were randomly assigned to the experimental condition. In the *hierarchy condition*, it was randomly determined which player will get 1/3 and wich player – 2/3 of the total payoff of the pair.

Subjects who have previously played the Prisoner's dilemma game were not allowed to participate in the study.

Results

To explore the influence of payoff distribution model on choices and cooperation in the PD games, the following dependent variables are analyzed: number of **cooperative choices for each player**; number of games with **mutual cooperation in a pair**; number of games with **mutual defection in a pair**. In the figures results are presented in percentages for clarity. However, the analysis is performed using the specified dependent variables.

Average **total payoff for a dyad (in points)** is analyzed to assess which type of payoff distribution led to higher profits.

Each dependent variable is analyzed in ANOVA with distribution model as between-subject factor with 4 levels (*unity* vs. *hierarchy* vs. *equality* vs. *proportionality*).

Cooperation

The cooperative choices (%) for each distribution type are presented in Figure 3. The analysis shows a significant influence of the distribution type on the number of cooperative moves ($F(3, 76) = 4.49$, $p = 0.006$).

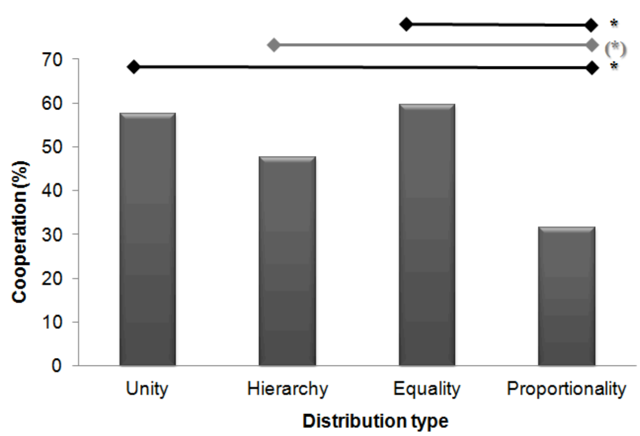

Figure 3: Average percentage of cooperative choices in each distribution condition ('*' means $p < 0.05$; '(*)' – marginally significant difference).

Post-hoc LSD test shows that the cooperation rate in the *proportionality condition* is significantly lower than the cooperation rate in the *unity condition* ($p = 0.003$) and in the *equality condition* ($p = 0.002$). The difference between cooperation rates in *proportionality condition* and *hierarchy condition* is marginally significant ($p = 0.065$). All other differences are non-significant.

It seems that the type of distributional model influences the cooperation rate. Cooperation is lower when each player gets a portion of the total payoff that is proportional to his/her payoffs during the game. In the terminology of the Fiske's theory, the market pricing relational model leads to diminished cooperation in comparison to the other three relational models. When the final payoff for the player depends on his/her individual results, the choices are more non-cooperative in comparison to the cases in which the total payoff of the pair is divided between players (no matter in what predefined proportions) and when each player received the total amount earned by both of them.

Mutual Cooperation

Average percentage of games in which there is mutual cooperation (both players have chosen to cooperate) is presented in Figure 4.

The ANOVA does not identify a statistically significant influence of the distribution type on the number of mutual cooperative game outcomes ($F(3, 36) = 1.94$, $p = 0.141$). However, a further conducted Post-hoc LSD test shows that difference exists between the *proportionality* and *equality condition* ($p = 0.038$). A marginally significant difference is observed between the *proportionality* and *unity condition* ($p = 0.079$).

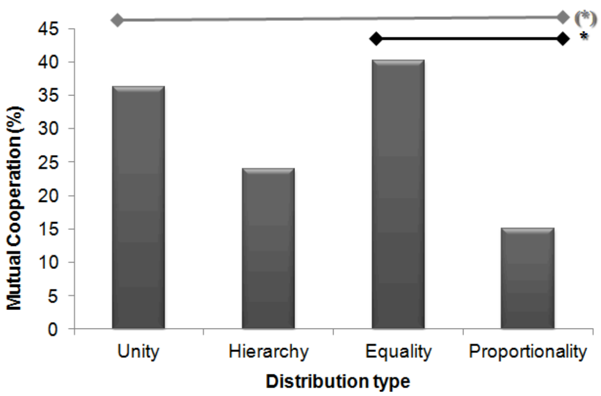

Figure 4: Average percentage of mutual cooperation in a pair in each distribution condition ('*' means $p < 0.05$; '(*)' – marginally significant difference).

It turns out that mutual cooperation is lower when the payoff of each player depends on his individual contribution compared to the situations in which players divide their joint winnings in equal shares or receive the total amount earned. This is a rather logical result in line with the assumption of the current study that the distribution type representing communal sharing or equality matching will promote collectivistic orientation; while money pricing relationship will most probably trigger an individualistic behavior among subject.

Mutual Defection

Average percentage of games with mutual defection (both players have chosen to defect) is presented in Figure 5.

For the number of games with mutual defection the ANOVA shows a significant influence of the distribution type ($F(3, 36) = 3.943$, $p = 0.016$). Post-hoc LSD test identifies differences between the *proportionality condition* and every other condition in the experiment – *unity* ($p = 0.006$), *hierarchy* ($p = 0.032$), and *equality* ($p = 0.005$).

Therefore, it can be concluded that when distribution of payoff is conducted according to individual results, mutual defection is a much more typical choice in comparison to all other distribution cases. In all other conditions this outcome (mutual defection) is relatively low. It should be noted that mutual defection leads to the lowest possible payoff for the pair. As noted in the introduction, the defection is dominant strategy for each individual player; however, mutual defection leads to the worst possible payoff for the society – thus the dilemma structure of the game arises as the opposition between individual and collective rationality.

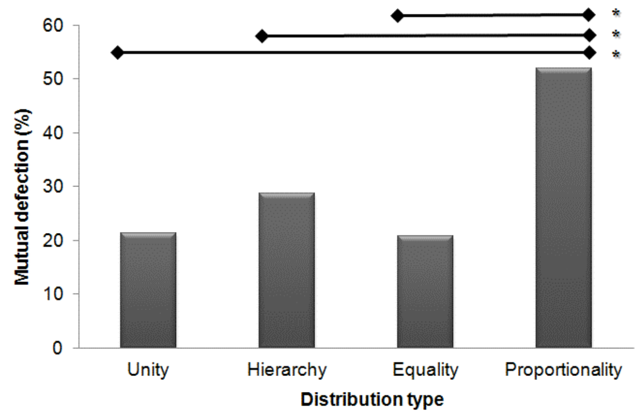

Figure 5: Average percentage of mutual defection in a pair in each distribution condition ('*' means $p < 0.05$).

Average Payoff

The payoff analysis was conducted on the basis of the average payoff per pair (in points) for the sequence of 40 games (Figure 6). The ANOVA shows a significant influence of the distribution type on the payoff ($F(3, 76) = 3.271$, $p = 0.026$).

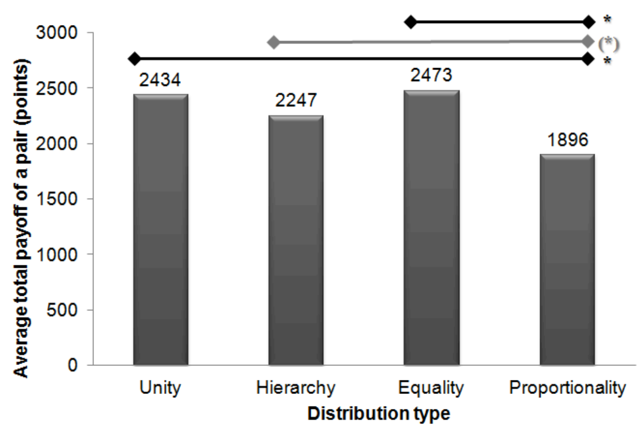

Figure 6: Average payoff per sequence of 40 games for a pair in each distribution condition ('*' means $p < 0.05$; '(*)' – marginally significant difference).

Significant differences were established through post-hoc LSD test between the *proportionality* and *unity condition* ($p = 0.011$) and between the *proportionality* and *equality condition* ($p = 0.006$). Marginally significant is the difference between the *proportionality* and *hierarchy condition* ($p = 0.093$).

The total payoff of a pair is lower when its distribution among the players depends on the individual contribution and points earned. This is a rather paradoxical result taking into account that such a distribution is representing the market pricing relational model, which is mostly related to individualistic attitude and profit orientation. However, taking into account that in the *proportionality condition* there is the highest number of games with mutual defection, the result is not surprising and could be explained by the lower payoff that the players get when both are non-cooperative.

Conclusions and Discussion

The current paper presents an experimental study on the phenomenon of cooperation in Prisoner's dilemma game in the light of Fiske's Relational models theory. The experiment was designed to explore the influence of the different distribution types of the total payoff (thus reflecting the four elementary human relations according to Fiske's theory) on the level of individual and mutual cooperative and non-cooperative behavior.

The results show that the distribution type corresponding to the individualistic Market Pricing relational model is characterized by a lower cooperation, lower mutual cooperation, higher mutual defection and lower total payoff of the participating subjects in comparison to the other distribution situations. When players are rewarded based on their individual results, they cooperate less and receive lower payoffs (for each player and for the dyad of players). This is an interesting result taking into account the fact that in formal game theory, in many experiments, in many real life situations, the players are perceived as individualistic beings. And there are attempts to apply policies aimed at achieving higher collective payoff by profit distribution accordingly to the contribution of each individual. Current results demonstrate that such distribution in fact leads to significantly lower total collective earnings.

An interesting topic for reflections and further research is whether a behavioral model representing competitiveness and profit-orientation might actually be effective in achieving its goals within the Prisoner's dilemma game model, respectfully within real life situations depicting this model.

Another related direction for exploration is *hierarchy condition*. Generally speaking, this condition could lead to higher cooperation despite the difference in received payoffs if the relation between the players is perceived as an in-team relation. This could be achieved by justifying the role attribution in this condition with some game related advantage (e.g. a better strategy in the test games).

A broader implications of these results might be applicable not only to studies of decision making in games, but also to socio-economical policies employed by organizations, governments, etc.

References

Andreoni, J., & Miller, J. (1993). Rational cooperation in the finitely repeated Prisoner's dilemma: experimental evidence. *The Economic Journal, 103*, 570-585.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390-1396.

Colman, A. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26,* 139-198.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: Experimental evidence from Prisoner's dilemma games. *Games and Economic Behavior, 12*,187-218.

Dawes, R. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169-193.

Durkin, J., Frost, M., Aronow, E., & Breslow, M. (1967). Moment of truth encounters in prisoner's dilemma. *American Psychologist*, *22*, 595.

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. New York, NY: Free Press.

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological Review*, *99*, 689-723.

Fiske, A. P., & Haslam, N., (1996). Social cognition is thinking about relationships. *Current directions in psychological science*, *5*, 143-148.

Gärdenfors, P. (2003). *How Homo became Sapiens*. Oxford University Press.

Grinberg, M., Hristova E., & Lalev, E. (2010) Models for cooperative decisions in prisoner's dilemma. In: Nefti, S. & Gray, J. (Eds.), *Advances in Cognitive Systems*. IET: London.

Haslam, N. (Ed.) (2004). *Relational models theory: a contemporary overview*. Mahwah, NJ: Erlbaum.

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

Kreps, D., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated Prisoner's dilemma'. *Journal of Economic Theory*, *27,* 245-252.

McGraw, A. P., Tetlock, P. E., & Kristel, O. V. (2003). The Limits of fungibility: relational schemata and the value of things. *Journal of Consumer Research, 36*, 219-229.

Rapoport, A., & Chammah, A. (1965). *Prisoner's dilemma: a study in conflict and cooperation*. Univ. of Michigan Press.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 117*, 57-75.

Sally, D. (1995). Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7,* 58-92.

Sally, D. (2001). On sympathy and games. *Journal of Economic Behavior & Organization, 44*, 1-30.

Trivers, R. (1972). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-37.