Mentalizing in games: A subtractive behavioral study of Prisoner's Dilemma

Antonio Napoli (antonio.napoli@phd.units.it)
Danilo Fum (Fum@units.it)
Dipartimento di Psicologia, Università degli Studi di Trieste
via S. Anastasio, 12, I-34134 Trieste, Italy

Abstract

Economists and neuroscientists often explain game playing by assuming that humans try to predict the opponent's behavior on the basis of her past choices. We try to question this assumption in a Prisoner's Dilemma Game by using a methodology which we call the "subtractive behavioral method". Our aim is to investigate which task features make participants attend to the opponent's behavior or, on the contrary, make them take into account only their own choices and received payoffs. We find a critical effect of contextual information and we derive some suggestions about the methodology of brain imaging and behavioral game theory experiments.

Keywords: Game Theory; Brain Imaging; Theory of Mind; Social Dilemmas; Prisoner's Dilemma

Introduction

Game Theory (Von Neumann & Morgenstern, 1944) is a branch of applied mathematics focused on describing and predicting the behavior of "players" involved in strategic interactions in which the result of every player's "move" is contingent on the move(s) made by the other player(s). One of the critical assumptions of the theory is that games are played by completely rational agents whose strategies could be precisely calculated. In recent years the Game Theory formalism has been adopted to develop models that try to account for the fact that people often behave differently from what the theory predicts. This approach has been named "Behavioral Game Theory" (Camerer, 2003).

Behavioral Game Theory models make the assumption that people learn during the interaction, i.e., that they change their behavior according to the efficacy of their past choices. Among these models there are some, like those based on Reinforcement Learning (Erev & Roth, 1998; Sarin & Vahid, 2001), which take into account only the player's own choices and received payoffs while others, like so-called sophisticated (Camerer, Ho, & Chong, 2002) and belief learning (Cheung & Friedman, 1997) models, consider also (or only) the opponent's choices and payoff history. We will refer to the former as "partial information models" and to the latter as "full information models".

Even if Behavioral Game Theory does not make any assumption about the internal mechanisms involved in game playing, from a cognitive perspective it is possible to find a difference between partial information and full information models. Partial information models obey to a strictly behaviorist rule: the more you get from a choice, the more you will choose it in subsequent trials. These models completely ignore the opponent's behavior and only manipulate representations about chosen moves and obtained payoffs. They may also be applied to situations of playing without opponents (one-person games); in fact, they have been proposed by Sutton and Barto (1998) to model the performance in multi-armed bandit tasks in which participants make repeated choices among different options which are followed by a numerical reward that depends on the choice being made.

On the other hand, full information models manipulate representations about the opponents' moves and payoffs to anticipate their behavior and obtain thus a strategic advantage. These models address the opponent's beliefs, intentions, and strategies, and therefore mimic a Theory of Mind (henceforth: ToM) or "mentalizing" mechanism.

Neuroscientist have recently begun to study the cortical circuits involved in game playing through neuroimaging. Krueger, Grafman, and McCabe (2008), after reviewing the literature on the topic, propose that two cognitive mechanisms are specifically involved in game playing.

The first one is a "shared affect system" located in the Anterior Insula. This area is only activated in non-zero sum games in which cooperation between players is possible, and therefore feelings of trust, reciprocity and collaboration could be developed. The area seems responsible of two main effects: it makes people feel disgust towards uncooperative behavior and react to it (for example, rejecting unfair offers in a Ultimatum Game: Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) and it makes people reciprocate by distinguishing between cooperative and non-cooperative opponents (Singer, Kiebel, Winston, Dolan, & Frith, 2004).

The second mechanism is a "shared intentions system", which is located in the Medial Prefrontal Cortex (MPFC). This area is activated both in zero and non-zero sum games, because it has the function of representing the opponent's beliefs, desires, and intentions, i.e. it seems to constitute the neural substrate of the ToM. Several brain imaging studies (see Krueger et al., 2008, for a comprehensive review) have shown MPFC activation during game playing and, therefore, it seems plausible that people mentalize while playing these games.

There are two other circuits which are not specifically involved in game playing but seem to be engaged in all kinds of learning tasks: a reward-based mechanism situated in a broad network of cortical and subcortical areas (see Lee, 2005 for a review), and a system concerned with the

prediction of complex behavior independently of its source, which is located in the Posterior Superior Temporal Sulcus (Frith & Frith, 2003).

Studies about mentalizing in game playing usually rely on the comparison between a condition in which people play against a computer and one in which they play against a human opponent on the presumption that mentalizing could be promoted by the latter. However, it is not clear whether and when people adopt a "mentalizing stance" and which task features could promote this activity. In fact, some studies show that a computer opponent could elicit activity in MPFC (Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004), while others claim that not all game situations against humans make people mentalize (Sally, 2003).

It is also unclear how mentalizing affects behavior, or, in other words, how decision making is affected by a ToM. For example Hill, Sally and Frith (2004) report that autistic adults behave in the same way as healthy participants in the Prisoner's Dilemma game, even if the autistic participants are severely impaired in other ToM tasks. Also, most neuroimaging studies lack a comparison between participant's behavior while playing against a human and a computer opponent.

We are convinced that the study of the ToM mechanisms would benefit from experiments which analyze participant's behavior. Two questions are important to us: 1) Which task feature make people mentalize? 2) Which effect does mentalizing have on people's behavior? In the present paper we try to address the first question by investigating some of the task features which could promote mentalizing during game playing.

Previous work

We have already started to explore the behavioral effects of mentalizing (Napoli & Fum, 2009) in playing a computer version of Rock, Papers, and Scissors (henceforth: RPS).

We had three groups of participants play 100 turns of RPS. In the first group, the computer was presented as an opponent, and the game was explicitly described as RPS. In the second group, the computer was presented as a neutral device. Participants saw three geometric figures which they should choose among at each trial; they received a payoff after each choice, and they could see the payoffs they could have obtained by making the alternative choices. Thus, this condition was equivalent to a multi-armed bandit task (Sutton & Barto, 1998) with the indication of foregone payoffs. In the third group, the computer was presented as an opponent. The game was played with the same rules of RPS but the choices were represented by geometric figures and the hierarchy of the moves (what beats what) had to be discovered during the game. This condition served as a control for the effect of the knowledge of the payoff matrix. The algorithm which assigned the payoffs was the same in all groups; the conditions differed therefore only for the setting induced in the participants (and the user interface).

We did not find any behavioral difference between the conditions, and we were able to model the behavior of all the groups by using a Reinforcement Learning algorithm based on ACT-R's utility learning mechanism (Anderson, 2007). This corroborates the idea that people did not use any information about foregone payoffs in the second condition and did not use any information about the opponent's moves or payoffs in the first and third condition. In summary, participants did not seem to mentalize at all during the experiment.

There are many possible explanations for this "failure to mentalize". Maybe people did not mentalize because they played against a computer; maybe they did not mentalize because the game was a mixed-strategy equilibrium game in which no move was better than the others and a simple behaviorist strategy could efficiently cope with the game; maybe people did not mentalize because no cooperation was possible in playing a competitive game. Or it may be a combination of all the three.

In this paper we try to clarify the findings of our previous work by making participants play a non-zero sum game, the Prisoner's Dilemma, both against what they believed was a human opponent and against a computer. Our aim is to understand which task features make people mentalize in game playing, which features affect game behavior and, possibly, why.

The experiment

Prisoner's Dilemma (henceforth: PD) is a non-zero sum game which has been extensively studied in psychology (Rapoport & Mowshowitz, 1966), classical game theory (Bo, 2005), behavioral game theory (Camerer, 2003), and neuroimaging studies (Singer et al., 2004). The payoff matrix used in our experiment is presented in table 1.

Table 1: Our experiment's payoff matrix

	Cooperate	**Defect**
Cooperate	60 60	100 0
Defect	0 100	20 20

PD can be thought of as a paradigmatic situation for any social dilemma in which the selfish interest contrasts with the common one. Classical game theory states that, independently of the choice made by the opponent, the most rational move for a player is to defect. In fact, if the opponent chooses to cooperate, defection gets 100 points and cooperation only 60 while, if the opponent defects, defection gets 20 points and cooperation 0. The result is that the optimal strategy for both people is to defect.

The most intriguing aspect of this game is that, even if the most rational move is defection, experiments show a substantial amount of cooperation between the players when the game is played in the iterated version (Bo, 2005). Another finding is that players learn to cooperate more and more during the experiment (Rapoport & Mowshowitz, 1966).

In order to understand what makes people mentalize, we adopted a "subtractive behavioral method" by assigning people to four different conditions in a repeated PD decision-making task in which the points earned by the participants were converted into play money.

The conditions differed according to the task features present in them which are summarized in Table 2.

Table 2: Features present in the experimental conditions

	Conditions			
Features	**N**	**CB**	**HB**	**HPD**
Repeated decision making	Y	Y	Y	Y
Opponent	N	Y	Y	Y
Believed Human Interaction	N	N	Y	Y
Explicit social scenario	N	N	N	Y

In the first condition, named "Nature" (N), participants played the PD disguised as a binary decision task: in each trial they had to choose between two options receiving a reward after each choice. It should be noted that in this condition the PD is presented as a repeated decision making one-person game, or a game against nature, in which no opponent is involved.

In the second condition, named "Computer Bet" (CB) participants were told that they would play a game against the computer. The instructions, however, presented the PD as a betting task: in each trial, the participants and the computer should bet on one of two alternatives and, depending on the combination of their choices, they would receive a given reward.

The third condition, named "Human Bet" (HB), was similar to the previous one (CB) except for the fact that participants were made to believe that they would play against a human opponent while in fact they were engaged by the computer.

In the fourth condition, named "Human Prisoner's Dilemma" (HPD), participants played PD against what they believed was a human opponent, just as in CB condition. There was, however, a substantial difference in the instructions provided for this condition and the two betting ones: the game was introduced by a story which illustrated a classical PD scenario (see Procedure for more details) and the two choices were labeled as "Cooperate" and "Defect".

In CB, HB and HPD conditions the instructions stressed that the goal of the participants was to gain as much money as possible independently of the money gained by the opponent, and that their opponent had the same objective.

According to results of neuroimaging research discussed in the Introduction, there are four cognitive processes which may influence participants' behavior in this task: the reward-based system, the complex behavior detecting system, the shared intentions system, and the shared affect system.

It is known that the reward-based system plays a role both in individual learning tasks and in game playing (Lee, 2005) by integrating the information received during the task in order to calculate the expected utility of different choices. Thus, this system should be active in all conditions, because of the repeated nature of the task.

It has been shown that the complex behavior detecting system is active during game playing against both computer and human opponents (Gallagher, Jack, Roepstorff, & Frith, 2002; Haruno & Kawato, 2009), and thus it should be activated in all conditions except Nature.

The shared intentions system is the main concern of this article. This area is always activated during game playing against humans, but it has been shown to be activated also during game playing against computer opponents, even if it is unclear which effect it exerts on people's behavior. If we find any difference between the CB and HB conditions, we can argue that mentalizing has a behavioral effect only in the case of a human opponent.

Finally, the shared affects system has been shown to be active when game playing involves the possibility of pro-social behavior, reciprocity, or fairness, and therefore we expect it could influence people's behavior only in the HPD condition. In this case the instructions promote empathizing with the opponent both because of the explicit social scenario and because of the labels attributed to the choices, which have a strong moral connotation. Therefore, every difference between the HB and HPD conditions should be attributed to this system.

Method

Participants and design. Sixty-four students (38 males) enrolled at the University of Trieste, Italy, were recruited as participants. Their age varied between 18 and 29 years (M=21.2, SD=3.4). Participants played two PD rounds, each one against a different algorithm (see below) whose order was counterbalanced between rounds. The experiment followed therefore a 4x2 mixed design with Setting as between-subjects and Algorithm as within-subjects factors.

Materials. Two algorithms were used in the experiment. The first one, Tit for Tat, cooperated in the first interaction and then replicated the opponent's previous choice. The second one, named Biased, made his moves by randomly sampling from a distribution of 60% Cooperate and 40% Defect moves.

Procedure. The experimental sessions were held in groups of 10-12 participants convened in a computer laboratory. Each participant was randomly assigned to one of the four conditions taking care that participants assigned to the same condition were not sitting next to each other. Participants were told that they would play different versions of the same game and received the instruction according to the condition to which they were assigned. Then, they were engaged in two PD rounds lasting eight minutes each.

The interface was kept as similar as possible in the four conditions. Participants made their choices by clicking on one of two circles displayed in the upper part of the screen. After a random lag time, in the Nature condition participants received a feedback about the money gained in the trial,

while in the other conditions they received a feedback about the opponent's choice, the money gained by themselves and by the opponent. The length of a bar representing their running total was updated and they were allowed to make another choice. In all conditions the two circles were labeled as "Yellow" and "Blue" except for the HPD condition, in which they were named as "Cooperate" and "Defect".

The main differences between the conditions relied in the amount of information and the kind of instructions provided to participants. In the N condition it was stated that they would play a binary decision task. After the first round participants were told that the computer would change the rule according to which it assigned the money. In the other three conditions participants had the payoff matrix in front of them from the beginning of the game. In the CB condition instructions stated that they would play a betting game with the computer, and after the first round they were told that the computer would change its strategy. In the HB and HPD condition participants were told they would play the game with one of the other participants in the room, and that the opponent would change after the first round. In the HB condition the task was presented as a betting game while in the HPD condition the game was introduced through a bargaining scenario in which Cooperate meant to respect the contract by delivering the promised goods and valuable money, respectively, while Defect meant to give the other player an empty bag. The instructions explicitly underlined this aspects of moral obligation and contract infringement involved in the game.

All groups played against the same algorithms with the Yellow and Blue circles equated to Defect and Cooperate, respectively.

At the end of the experiment we had informal interviews with the participants to assess the possibility that they had some doubts about having played against a computer and not a mate. Subjects who reported doubts were discarded from data analysis. Finally, a collective debriefing session ensued in which the nature of the opponent was discovered to all participants and the reasons for always adopting a computer as opponent were explained.

Results

Since the experiment was self-paced, participants made a variable number of choices in each round. To perform statistic analyses, we took into account their first 50 moves only.

Analysis of Cooperations. Being interested in the quality of participant's behavior more than in their ability to exploit the opponent's algorithm, we concentrated the analysis on the number of Cooperate moves and not on the amount of money gained.

First, we looked for possible differences between the first and second round in order to control for effects of learning (or fatigue). A mixed design ANOVA between the Round and the Setting did not reveal any significant effect for the Round (p=.55) or interaction (p=.93), while there was a significant effect of the Setting (F(3,58)=10.1, p < .001).

We then analyzed the factors manipulated in the experiment. A mixed design ANOVA revealed a significant effect of Setting and Algorithm (F(3,58)=10.1, p<.001 and F(1,58)=93.14, p<.001 respectively), while the interaction was not significant (p=.92). Table 3 reports Means and Standard Deviations of the participants' total Cooperate moves.

Table 3: Means (and Standard Deviations) of Cooperate per Algorithm in the various Settings

	Setting			
Algorithm	N	CB	HB	HPD
Biased	**15.69** (5.41)	**11.18** (7.7)	**14.23** (10.03)	**24.24** (7.09)
TFT	**32.6** (9.93)	**26** (10.65)	**28.8** (17.2)	**41.35** (10.6)

Algorithm and Setting seem to have an addictive effect in promoting cooperation between participants. While it is evident that the TFT algorithm promotes Cooperation more than the Biased one, it is unclear how Settings exerted its effect. Since there was no main effect of Round and no interaction between Algorithm and Setting, we analyzed separately the participant's performance against the two algorithms.

Two separate one-way ANOVAs for Biased and TFT Algorithms were performed. Both showed a significant effect for Setting (F(3,58)=8.95, p<.001 and F(3,58)=4.94, p<.01 respectively). The probabilities associated with post-hoc Newman-Keuls tests to contrast each Setting condition with the others are summarized in tables 4 and 5. For both algorithms a significant difference was found between the HPD and the other three conditions which, on the other hand, did not differ from each other.

Table 4: Probabilities for Post-hoc Newman-Keuls tests for the Biased Algorithm

	N	**CB**	**HB**	**HPD**
N		.24	.6	.0029*
CB	.24		.27	.002*
HB	.6	.27		.0017*
HPD	.0029*	.002*	.0017*	

* = significant

Table 5: Probabilities for Post-hoc Newman-Keuls tests for the TFT Algorithm

	N	**CB**	**HB**	**HPD**
N		.29	.39	.051**
CB	.29		.52	.016*
HB	.39	.52		.0049*
HPD	.051**	.0049*	.016*	

*= significant **=marginally significant

Analysis of conditional probabilities. We ran another analysis in order to understand why there was a difference in the number of Cooperate moves in HPD condition. This analysis was proposed by Rapoport and Mowshowitz (1966) and was also utilized by Erev and Roth (2001) in order to assess the efficacy of their reinforcement learning model.

Rapoport and Mowshowitz analyzed the probability of cooperation in a given trial according to the choices made in the *previous* trial by *both* players. Thus, a participant's strategy can be described by four numbers, C|CC, C|CD, C|DC, and C|DD. In the N condition, these probabilities may be interpreted as an analysis of a "win stay / lose switch" behavior. We can assume that, after a few choices, people get acquainted with the payoffs associated with the various options. Thus, for example, C|CC would be the probability of making the Cooperate/Blue move after receiving the best reward associated with that choice; therefore, a high value of this parameter would be an expression of a "win stay" strategy.

We analyzed the four conditional probabilities separately for the two algorithms to search for possible different strategies used in the different Settings. We ran a total of eight one-way ANOVAs analysis and all post-hoc Newman-Keuls tests for the significant ones.

We found a significant difference in three ANOVAs: C|DC both in the Biased ($F(3,58)=7.94$, $p<.001$) and in the TFT condition ($F(3,58)=5.21$, $p<.005$) and C|CC in the TFT condition ($F(3,54)=4.73$, $p<.006$). Newman-Keuls post-hoc tests showed that: in C|DC / Biased, HPD was different from all the other conditions ($p<.001$ in all cases), which were similar between them; in C|DC / TFT, HPD was different from CB and HB ($p<.001$ in both cases) and only marginally significant respect to N ($p=.055$), and the other three conditions were similar between them; in C|CC / TFT, the only significant difference was between HPD and CB $p<.001$.

Discussion and conclusions

In the experiment, participants played against an algorithm, the Biased one, that chooses its moves by sampling randomly from a given distribution, i.e., independently from the move made by the opponent, and against another algorithm, the TFT, that cooperates only if the opponent cooperated in the previous trial and defects otherwise. This means that the most rewarding strategy for participants was to Defect against the Biased algorithm—in order to exploit the trials in which it cooperates and to defend against the possibility of being exploited when the algorithm defects—and to Cooperate against the TFT—in order to initiate and maintain a virtuous reciprocation loop. The statistical analyses demonstrated that participants made more Cooperate moves against the TFT than against the Biased algorithm, i.e., that they were successful in adapting their strategy to the strategy used by the opponent.

However, we also found some differences between the groups: in the HPD condition participants made a higher number of Cooperate moves against both algorithms. The conditional probability analysis showed that this difference could be explained by the higher rate of C|DC in both cases. Since the only difference between the HPD and the other groups relied in the use, in the former case, of instructions that explicitly underlined the aspects of moral obligation and contract infringement involved in the game, the most natural conclusion is that this feature made people more prone to regret their defection against a cooperative opponent in the previous trial leading thus to more frequent cooperative behavior.

Interpreting the behavioral results in terms of the cognitive systems framework introduced above, we could safely assume an influence on this task of the reward-based system, being the participants capable of successfully adapting their strategy to the opponent in all conditions. However, we cannot exclude that such a performance could reflect the activation of the complex behavior prediction system too, being the activation of this system not selectively associated with strategic interactions (Frith & Frith, 2003). As for the shared affect system, it could have played a role in both human conditions (HB and HPD). In fact, during the debriefing interviews, some HB participants spontaneously told us about their willingness to cooperate with the opponent, a behavior that is typically associated with the activation of this system (Singer et al., 2004). However it is unlikely that this system played a critical role in the HB group, whose performance was similar to that of the N and CB condition where it is not credible that people could empathize with a computer, being it an opponent or not. Therefore, this system could be active only in the HPD condition.

As for the ToM system, we can exclude that it influenced the participant's behavior in CB and HB groups, which was similar to that of the N group. Therefore, we are left with two systems (ToM and empathizing) as responsible for the difference found in the HPD condition. Because brain imaging studies show that playing against a human opponent activates ToM areas regardless of the specific game (see for example Gallagher et al., 2002) and because, according to the participant's reports, it seems likely that they did in fact mentalize, we think that this area was active in both situations, and suggest two possible explanations for our results: (1) ToM had no behavioral effect in HB situation or (2) ToM had no effect both in the CB and HB conditions, and the difference between the two groups should be attributed to the shared affect system.

We won't take position with regard to this issue, because the limitations of our behavioral method don't permit us to. However we think that, whichever is the real explanation, this study makes some interesting points about both brain imaging and behavioral game theory experiments.

With regard to brain imaging studies, even if it has been shown that ToM areas are active in almost every game played against human opponents, it is not clear when they have a behavioral effect, too. We can speculate that there is some mechanism which prevents ToM from influencing the behavior in some situations. Otherwise, it would seem really

strange that it wouldn't have any effect on behavior *at all*. Therefore, we think that brain imaging studies should always take in account people's behavior, in a similar way to Haruno & Kawato (2009) and Hampton, Bossaerts, and O'Doherty (2008).

As for behavioral game theory, this paper makes a case for Erev and Roth's (2001) proposal of accounting people's behavior in Prisoner's Dilemma by the means of Reinforcement Learning. In fact, in N condition participants did not have any information about foregone payoffs, and nonetheless, their behavior was similar to the other groups. This means that the knowledge of payoff matrix and of the opponent's choices had a limited effect on participant's behavior. On the other side, the paper shows also the importance of contextual information—a variable which is seldom taken into account in game theory. In a more general sense, we think that our paper suggests the utility of having, along experiments in which people play one against the other, some more controlled sessions in which the participants play against an opponent (be it a computer or a human actor) whose strategy was under the control of the experimenter and compare them with individual learning sessions. This could make the experimenter safely exclude in most cases unnecessary believes or sophisticated learning.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Bo, P. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review, 95,* 1591–1604.

Camerer, C. F. (2003). *Behavioral game theory: Experiments on strategic interaction.* Princeton: Princeton University Press.

Camerer, C. F., Ho, T., & Chong, K. (2002). Sophisticated Experience-Weighted Attraction learning and strategic teaching in repeated games. *Journal of Economic Theory, 104,* 137–8.

Cheung, Y., & Friedman, D. (1997). Individual learning in normal form games: some laboratory results. *Games and Economic Behavior, 19,* 46–76.

Erev, I. & Roth, A. (1998). Predicting how people play games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review, 88,* 848–881.

Erev, I. & Roth, A.E. (2001). On simple reinforcement learning models and reciprocation in the prisoner dilemma game. In Gigerenzer, G. and Selten, R. (Eds.), *The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society, 358*, 459–473.

Gallagher, H.L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the Intentional Stance in a Competitive Game. *Neuroimage, 16,* 814–821.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Procedings of the National Academy of Sciences USA, 105*, 6741 – 6746.

Haruno, M., & Kawato, M. (2009). Activity in the Superior Temporal Sulcus Highlights Learning Competence in an Interaction Game. *The Journal of Neuroscience, 29,* 4542–4547.

Hill, E. L., Sally, D. & Frith, U. (2004). Does mentalizing ability influence cooperative decision – making in a social dilemma? *Journal of Consciousness Studies, 11,* 144 – 161.

Krueger, F., Grafman, J., & McCabe, K. (2008). Neural correlates of economic game playing. *Philosophical Transactions of the Royal Society, 363*, 3859 – 3874.

Lee, D. (2005). Neural basis of quasi-rational decision-making. *Current Opinion in Neurobiology, 16,* 191–198.

Napoli, A., & Fum, D. (2009). Applying Occam's razor to paper (and rock and scissors, too): Why simpler models are sometimes better. *Proceedings of the 9th International Conference of Cognitive Modeling*, Manchester, United Kingdom.

Rapoport, A., & Mowshowitz, A. (1966). Experimental studies of stochastic models for the prisoner's dilemma. *Behavioral Science, 11,* 444–458.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage, 22,*1694–1703.

Sally, D. (2003). Dressing the mind properly for the game. *Philosophical Transactions of the Royal Society of London B, 358,* 583 – 592.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. 2003 The neural basis of economic decision-making in the Ultimatum game. *Science, 300*, 1755–1758.

Sarin, R. & Vahid, F. (2001). Predicting how people play games: a simple dynamic model of choice. *Games and economic behavior, 34,* 104–22.

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J. & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron, 41*, 653–662.

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

Von Neumann, J., & Morgenstern, O. (1944) *Theory of games and economic behaviour.* Princeton, N.J.: Princeton University Press.