A Computer Model of the Interpersonal Effect of Emotion Displayed in a Social Dilemma*

Celso M. de Melo[1], Peter Carnevale[2], Dimitrios Antos[3], Jonathan Gratch[1]

[1] Institute for Creative Technologies, University of Southern California, 12015 Waterfront Drive, Building #4 Playa Vista, CA 90094-2536, USA
{demelo, gratch}@ict.usc.edu
[2] USC Marshall School of Business, Los Angeles, CA 90089-0808, USA
peter.carnevale@marshall.usc.edu
[3] Harvard University, 33 Oxford st., Maxwell-Dworkin 217, Cambridge, MA 02138, USA
antos@fas.harvard.edu

Abstract. The paper presents a computational model for decision-making in a social dilemma that takes into account the other party's emotion displays. The model is based on data collected in a series of recent studies where participants play the iterated prisoner's dilemma with agents that, even though following the same action strategy, show different emotion displays according to how the game unfolds. We collapse data from all these studies and fit, using maximum likelihood estimation, probabilistic models that predict likelihood of cooperation in the next round given different features. Model 1 predicts based on round outcome alone. Model 2 predicts based on outcome and emotion displays. Model 3 also predicts based on outcome and emotion but, considers contrast effects found in the empirical studies regarding the order with which participants play cooperators and non-cooperators. To evaluate the models, we replicate the original studies but, substitute the humans for the models. The results reveal that Model 3 best replicates human behavior in the original studies and Model 1 does the worst. The results, first, emphasize recent research about the importance of nonverbal cues in social dilemmas and, second, reinforce that people attend to contrast effects in their decision-making. Theoretically, the model provides further insight into how people behave in social dilemmas. Pragmatically, the model could be used to drive an agent that is engaged in a social dilemma with a human (or another agent).

Keywords: Emotion, Cooperation, Social Dilemma, Probabilistic Model

1 Introduction

In multi-agent systems, agents frequently have to decide whether to pursue their own self-interest and collect a short-term reward or trust other agents to reach mutual cooperation and maximize joint long-term reward [1]. Initial solutions to such

* This research supported by the Air Force Office of Scientific Research under grant FA9550-09-1-0507 and the National Science Foundation under grant IIS-0916858. The content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

dilemmas were based on game-theoretic notions such as dominant strategies or Nash equilibria that prescribe the conditions under which it is rational to cooperate [2]. However, though appropriate for agent-agent encounters, these techniques are less so for human-agent encounters. Effectively, there is now considerable evidence that humans are not purely self-interested and do not always behave according to the predictions of game theory [3, 4]. Early research in the behavioral sciences has, in fact, shown many sources of cooperation in human-human interaction [5]: some people are simply inclined to cooperate [6]; group identity [7]; reciprocity [8]; monitoring and sanctioning [9]; and, verbal communication [10]. More recently, non-verbal displays have also been argued to impact emergence of cooperation [11, 12], in particular, facial displays of emotion (e.g., [13]).

In a pioneering set of studies [14-17], we have explored the interpersonal effect of emotion displays on emergence of cooperation between agents and humans in social dilemmas. In these studies participants play the iterated prisoner's dilemma with agents that, even though following the same strategy to choose their actions, convey different facial emotion displays according to the outcome of each round. In line with predictions from the behavioral sciences about the impact of non-verbal displays on decision-making [11, 12], the results indicate that people's decision to cooperate is influenced by emotion displays. For instance, people cooperate more with an agent which displays reflect an appreciation of cooperation (e.g., smile when both players cooperate) than one which displays reflect satisfaction with selfishness (e.g., smile when agent defects and participant cooperates). In line with the view that people respond emotionally to *relative* changes in their situations rather than the absolute consequences of their decisions [4], the results show that the order participants play the agents influences cooperation rates. For instance, people will cooperate more with an agent which displays reward cooperation *after* playing with one which displays reflect selfish interest, than if they were to play with the former agent first.

In this paper we develop a computer model of decision-making in a social dilemma that takes into account the other party's emotion displays and replicates findings from the literature on how people behave in social dilemmas. Such a model could be used to drive embodied agents - i.e., agents that have virtual bodies and can express through them like humans do [18] - when engaged in a social dilemma with humans. Effectively, it has been shown that people can treat embodied agents like people [19] and are capable of being influenced by them [20]. Moreover, since embodied agents can be used to learn about human-human interaction [21], such a model would allow us to get further insight on how people act in social dilemmas. Methodologically, we follow a novel approach: (1) Data from our empirical studies [14-17] is collapsed into a single database. Features represent aspects of the game, the outcome of the round and whether the participant cooperated in the next round (target); (2) Probabilistic models are fitted to the data using maximum likelihood estimation. Each model predicts likelihood of cooperation given a subset of the features (e.g., outcome and display in the current round). We explore models that predict based on outcome only, outcome and emotion, and outcome, emotion and contrast effects; (3) Regarding evaluation, even though we look at standard performance measures such as error rate, the focus is on the models' ability to replicate previous findings about how people behave in social dilemmas. To accomplish this, we "play" the models with different agents that display emotions, under the same configurations as in the empirical studies. Our results show that the best model replicates many of the findings about how people behave in social dilemmas and, overall, reinforce findings for the importance of attending to nonverbal signals and contrast effects in social dilemmas.

2 Background

This section describes three empirical studies we previously conducted where people are engaged in a social dilemma with agents that display emotions through the face.

Study 1. The first study [15, 16] follows a repeated-measures design where participants play 25 rounds of the iterated prisoner's dilemma with two agents that play the same strategy but show different emotion displays. The prisoner's dilemma game was recast as an investment game where participants can choose to invest either in Project Green (cooperation) or Project Blue (defection). The payoff matrix is shown in Table 1. The agents' action strategy is based on tit-for-tat [8]. The *expressively cooperative* agent displays reflect an appreciation of mutual cooperation (e.g., when both players cooperate it smiles). In line with the definition of selfish orientation [6], the *expressively individualistic* agent's displays reflect how valuable the outcome is to the agent, independently of the value to the participant (e.g., when the agent cooperates and the participant defects, it shows sadness). Table 2 summarizes the displays for both agents. Agent order was counter-balanced across participants. Fifty-one participants were recruited for this experiment.

Table 1. Payoff matrix for the social dilemma game.

		Agent			
		Project Green		Project Blue	
Participant	Project Green	Agent:	5 pts	Agent:	7 pts
		Participant:	5 pts	Participant:	3 pts
	Project Blue	Agent:	3 pts	Agent:	4 pts
		Participant:	7 pts	Participant:	4 pts

Table 2. Emotion displays for the agents in study 1.

Expressively Cooperative		*Agent* Green	*Agent* Blue
Participant	Green	Joy	Shame
	Blue	Anger	Sadness

Expressively Individualistic		*Agent* Green	*Agent* Blue
Participant	Green	Neutral	Joy
	Blue	Sadness	Sadness

The results show that, as predicted, people's decision making is influenced by the emotion displays and people cooperate significantly more with the cooperative[†] agent. Additionally, the results reveal clear contrast effects: people cooperate more with the cooperative agent after playing with the individualistic agent, than the other way around. This contrast effect is in line with the well-known *black-hat/white-hat* (or *bad-cop/good-cop*) effect [22] that argues people cooperate more with a cooperative

[†] When the context is clear we refer to the agents without the 'expressively' adverb.

opponent if they're first matched with a tough opponent. In summary, the findings in this study are (see cooperation rates in Table 8 under 'Humans'):

F1.1 (a) *Participants cooperate significantly more with the cooperative than the individualistic agent;* ***(b)*** *but, this effect is mainly driven by the order where participants play with the individualistic agent first, followed by the cooperative agent.*

Study 2. The second study (unpublished) explores two new versions of the cooperative and individualistic agents that display the same type and quantity of emotions but, the displays are mapped differently to round outcomes. Table 3 summarizes these new agents. In this study we also compare the emotional agents to a no-emotion control agent. Three experiments were run: (1) cooperative vs. individualistic, with 39 participants; (2) cooperative vs. control, with 20 participants; (3) individualistic vs. control, with 37 participants. Otherwise, the design remained the same as study 1.

Table 3. Emotion displays for the agents in study 2.

Expressively Cooperative		*Agent* Green	*Agent* Blue
Participant	Green	Joy	Sadness
Participant	Blue	Sadness	Sadness

Expressively Individualistic		*Agent* Green	*Agent* Blue
Participant	Green	Sadness	Joy
Participant	Blue	Sadness	Sadness

The results show the following (see cooperation rates in Table 8 under 'Humans'):

F2.1 *Participants tend to cooperate more with the cooperative agent than the individualistic agent, in all orders;*

F2.2 ***(a)*** *Participants tend to cooperate more with the cooperative than the control agent;* ***(b)*** *but, this effect is driven by the order where participants play with the control agent first;*

F2.3 *Participants do not cooperate differently with the individualistic and control agents.*

Study 3. The third study [17] compares (a variant of) the expressively cooperative agent with the *expressively competitive* agent. In line with the usual definition of competitive orientation [6], the competitive agent's displays reflect a goal of earning more points than the participant (e.g., when the agent defects and the participant cooperates, it smiles). Table 4 shows the emotion displays for these agents. In this study, we also compare the emotional agents to a no-emotion control agent. We ran 3 experiments: (1) cooperative vs. competitive, with 34 participants; (2) cooperative vs. control, with 38 participants; (3) individualistic vs. control, with 30 participants. The design remains the same as study 1, except that the payoff for the player that gets exploited (i.e., when it cooperates and the other defects) is reduced from 3 to 2 points.

Table 4. Emotion displays for the agents in study 3.

Expressively Cooperative		*Agent* Green	*Agent* Blue
Participant	Green	Joy	Neutral
Participant	Blue	Anger	Neutral

Expressively Competitive		*Agent* Green	*Agent* Blue
Participant	Green	Neutral	Joy
Participant	Blue	Anger	Neutral

The results show the following (see cooperation rates in Table 8 under 'Humans'):

F3.1 ***(a)*** *When collapsing across orders, participants tend to cooperate more with the cooperative agent; however:* ***(b)*** *in the order cooperative → competitive, participants tend to cooperate more with the competitive agent;* ***(c)*** *in the order competitive → cooperative, participants cooperate significantly more with the cooperative agent;*

F3.2 ***(a)*** *Participants cooperate significantly more with the cooperative agent than the control agent;* ***(b)*** *but, this effect is mainly driven by the order where participants play with the control agent first;*

F3.3 *Participants do not cooperate differently with the competitive and control agents.*

3 Models

To develop a model for decision-making in social dilemmas, we follow a data-driven approach based on data collected in the aforementioned empirical studies.

Data and Features. The data consists of examples corresponding to each round each participant played in each study. Data corresponding to last rounds is ignored, since the goal is to predict whether the participant cooperates in the *next* round. In total, there are 12,432 examples. The feature set is the following:
(a) Outcome of the Round: whether the players cooperated or defected;
(b) Emotion Display: the agent's display following the outcome in that round;
(c) First Game: whether the example corresponds to a round in the first game;
(d) Agent is cooperator: 'true' if current agent is a cooperator;
(e) Previous Agent is Cooperator: 'true' if (eventual) previous agent is a cooperator;
(f) Whether Participant Cooperates in the Next Round: this is the target attribute.

Training, Validation and Test Sets. The data is first partitioned into a training (75%) and a test set (25%). The training set is further partitioned into 20 subsets to support 20-fold cross-validation. Every subset (including the test set) are created while making sure they have the same proportion of positive and negative examples in each of the three studies as in the whole dataset.

Models. Models consist of rules defining the probability of cooperation in the next round, given a subset of the features. We explore three different models, described below, that use different subsets of the features. Maximum likelihood estimation is used to fit the models to the data and estimate parameters. The training procedure does 20-fold cross-validation and the final model parameters correspond to the average over all training sets.

Model based on Outcome. The first model predicts likelihood of cooperation based only on outcome of the current round. Outcome is chosen as the first attribute as it ranks best according to the information gain metric (or Kullback–Leibler divergence). Thus, the model predicts probability of cooperation in the next round, given a certain outcome in this round. These probabilities are obtained by calculating the frequency the participant cooperated after each round, for each possible outcome. Table 5 shows the parameters (averaged over all training sets) for this model (under 'Model 1').

Model based on Outcome and Emotion Displays. The next model predicts likelihood of cooperation given outcome *and* the agent's display. This model's parameters are shown in Table 5 (under 'Model 2').

Model based on Outcome, Emotion Displays and Contrasts. Finally, the third model also tries to predict likelihood of cooperation based on outcome and emotion displays but, also takes into account the black-hat/white-hat contrast effects reported in our studies (see 'Background'). All the information required to represent these effects is in attributes (c), (d) and (e), i.e., attributes regarding whether the first and second agents are black-hats (non-cooperators) or white-hat (cooperators). However, these attributes are conceptually different than outcome and emotion displays, because they are *non-observable*. Effectively, they represent *inferences* participants make while playing the games. Nevertheless, notice these inferences *are* made, consciously or not, because otherwise there would have been no contrast effects. Still, for the time being, we do not attempt to model the mechanism by which participants make these inferences and simply assume that attributes (c), (d) and (e) are directly observable (but, see the 'Discussion' section for a way to address this in the future). In summary, the third model calculates, for each combination of the attributes (c), (d) and (e), probabilities given the outcome and agent's displays in the previous round (see Table 5 under 'Model 3'). Notice there is no prediction for the case where both the 1st and 2nd agents are white-hats because this was not explored in our studies.

Table 5. Parameters for the probabilistic models. Values represent probability of cooperation.

Outcome	Emotion	Model 1	Model 2	Model 3				
				BH *1st Game*	WH *1st Game*	BH→WH *2nd Game*	WH→BH *2nd Game*	BH→BH *2nd Game*
CC	joy	.67	.72		.64	.76		
	neutral		.62	.53			.71	.51
	sadness		.61	.61			.57	.54
DD	neutral	.22	.24	.22	.27	.30	.21	.26
	sadness		.20	.21	.20	.20	.17	.25
huCagD	joy	.29	.26	.30			.27	.16
	neutral		.26	.26	.15	.40	.19	.23
	sadness		.34		.35	.33		
	shame		.36		.27	.40		
huDagC	anger	.28	.27	.27	.28	.27	.37	.23
	neutral		.24	.22			.34	.19
	sadness		.31	.27	.30	.29	.34	.37

CC - mutual cooperation; *DD* - mutual defection; *huCagD* - human cooperates, agent defects; *huDagC* - human defects, agent cooperates; *BH* - Black-Hat (or non-cooperator); *WH* - White-Hat (or cooperator); *1st Game* refers to probabilities in the 1st game (with a BH or WH); *2nd Game* refers to probabilities in the second game (with a BH or WH) but, when the game was preceded by a specific first game (with another BH or WH)

Model Selection. Model selection is based on minimizing *error rate*, i.e., the percentage of incorrectly classified examples (averaged over all 20 validation sets). Table 6 shows the error rates for each model. The results show that error rates are significantly different ($F(2, 57)=28.207$, $p<.05$) and, *LSD* post-hoc tests reveal that:

the error rate for Model 1 is higher than for Model 2 (p=.100); and, the error rate for Model 2 is higher than for Model 3 (p=.000). Table 6 also reports several other standard measures (precision, recall, F1, etc.) and it is clear that Model 3 outperforms Model 2 which, in turn, outperforms Model 1. Table 7 reports the results over the test set. Error rate suggests, once again, that Model 3 is better than Model 2 and, in turn, Model 2 is better than Model 1. The remaining variables in Table 7 also generally support that Model 3 fares best and that Model 1 fares worst. Finally, *average log likelihood* measures the posterior probability of the (whole) dataset given the model, averaged over the number of examples (the closer to 0, the better). The results for the models are: Model 1, *-0.247*; Model 2, *-0.246*; and, Model 3, *-0.245*. Thus, the results suggest that the data was most likely to have been generated from Model 3 than any of the other models.

Table 6. Performance measures over validation sets.

	Model 1		Model 2		Model 3			
	Mean	SD	Mean	SD	Mean	SD	F	Sig.
error	.382	.016	.373	.017	.345	.017	28.207	.000*
precision	.408	.024	.422	.025	.466	.025	29.842	.000*
recall	.407	.025	.423	.026	.466	.024	29.571	.000*
F1	.408	.025	.423	.026	.466	.024	29.575	.000*
true positives	61.332	4.935	63.717	5.298	70.134	4.958	16.147	.000*
false positives	88.885	4.434	87.196	4.582	80.339	4.785	19.342	.000*
true negatives	226.566	3.599	228.254	3.760	235.112	3.886	29.136	.000*
false negatives	89.069	4.618	86.684	4.526	80.267	4.583	19.798	.000*

* significant difference, p<.05

Table 7. Performance measures over the test set.

model	error	precision	recall	F1	tp	fp	tn	fn
Model 1	.382	.411	.421	.416	422.86	606.99	1498.01	581.14
Model 2	.38	.414	.425	.419	426.72	605.25	1499.75	577.28
Model 3	.378	.417	.424	.421	425.85	595.55	1509.45	578.15

tp - true positives; fp - false positives; tn - true negatives; fn - false negatives

4 Evaluation

The results in the previous section suggest Model 3 was best and Model 1 worst at predicting how humans behave in these dilemma situations. However, in this section we explicitly test this by replicating our empirical studies [15-17] but, substituting humans for our probabilistic models. Aside from verifying the results from the previous section, this experiment allowed us to get insight into the mechanisms that explain *why* some models do better than others. To accomplish this, we ran each model 1000 times (500 times per order) for each experiment in our studies, and measured which findings (F1.1 to F3.3, see “Background”) the models replicate. The cooperation rates and standard deviations for the original human data and the models are shown in Table 8. Two columns are shaded in this table, for each model: (1) the left column summarizes whether cooperation rates were significantly different (p<.05) and represent an effect size above a minimum threshold[‡], which we set to 1.5 (corresponding to, at least, a small effect size). For instance, a '>' means the model cooperated significantly more with the agent on the left than the agent on the right and the effect size passed the threshold; (2) the right column shows a tick if the model successfully replicated the findings in the human data. Therefore, the more ticks a model has, the better it is at replicating findings. Overall, the percentage of findings each model replicated was: Model 1, *42.9%* (9 out of a maximum of 21 ticks); Model 2, *81.0%* (17 out of 21 ticks); and, Model 3, *95.2%* (20 out of 21 ticks).

5 Discussion

In this paper we propose a data-driven probabilistic model for decision-making in a social dilemma when the other party displays emotion. The evaluation reveals that the model is better at replicating findings about how humans behave in social dilemmas if, instead of considering round outcome alone, it also considers emotion displays. This result is in line with predictions in the behavioral sciences about the impact of non-verbal displays on decision-making [11, 12]. The results also show that considering (black-hat/white-hat) contrast effects further improves the ability to predict human behavior. This is in line with the view that people respond emotionally to *relative* changes in their situations rather than the absolute consequences of their decisions [4]. Theoretically, the model complements the findings in our original studies [14-17] by quantizing (through probabilities) the effect of emotion displays on decision-making in a social dilemma. For instance, Model 2 (see Table 5) suggests that, after the human is exploited by the agent (i.e., when the human cooperates and the agent defects), the human's likelihood of cooperating goes up from 26% to 36% if the agent displays shame as opposed to joy. Finally, pragmatically, the model can be used to drive an agent that is engaged in a social dilemma with another human (or agent) that shows emotion.

There is, naturally, much future work ahead: (1) error rates (Table 6 and 7) are still relatively high and this might reflect that important features that characterize how people decide in social dilemmas are being neglected. For instance, it is assumed that examples are independent and identically-distributed (i.i.d.), but this is not strictly accurate (e.g., people tend to defect towards the end independently of the agent they're playing with); (2) model 3 assumes it is known whether the other party is a black- or white-hat but, in fact, this information should be inferred. One way to address this is to use a Bayesian learning mechanism that increases the likelihood of the opponent being a black-hat according to the displays it shows for each outcome; (3) there are combinations of outcome and displays for which there are no examples in the database. To address this we need to run new experiments where participants face agents with the missing combinations of outcome and display; finally, (4) to further test the generalizability of the model, a new sample should be gathered with human participants and the results compared to the model's predictions.

---

‡ Because it's possible to get significance even for small differences if the sample size is large enough, it is important to require the effect size to be above a minimum threshold.

Table 8. Evaluation of the probabilistic models. Cooperation rates (standard deviations) are shown for the original empirical data (under 'Humans') and when running the models under each of the experimental configurations. The left-most shaded column summarizes the comparison between cooperation rates between the two agents in that configuration. The right-most shaded column is interpreted as follows: ✓ means the model replicates the findings in the human data; × means the model doesn't replicate the human data.

		Humans				**Model 1**				**Model 2**				**Model 3**			
Study 1	*Order*	*Cooperative*	*Individual.*			*Cooperative*	*Individual.*			*Cooperative*	*Individual.*			*Cooperative*	*Individual.*		
Coop	both	.37 (.28)	.27 (.23)	>	*F1.1a*	.33 (.14)	.33 (.14)	≈	×	.36 (.16)	.31 (.13)	>	✓	.35 (.16)	.31 (.14)	>	✓
Vs.	coop→indiv	.35 (.26)	.31 (.26)	≈	*F1.1b*	.32 (.14)	.33 (.14)	≈	✓	.37 (.17)	.30 (.12)	>	×	.31 (.14)	.32 (.15)	≈	✓
Indiv	indiv→coop	.39 (.30)	.23 (.19)	>	*F1.1b*	.33 (.14)	.33 (.14)	≈	×	.35 (.16)	.31 (.13)	>	✓	.39 (.17)	.30 (.12)	>	✓
Study 2		*Cooperative*	*Individual.*			*Cooperative*	*Individual.*			*Cooperative*	*Individual.*			*Cooperative*	*Individual.*		
Coop	both	.39 (.24)	.33 (.24)	>	*F2.1*	.33 (.14)	.33 (.14)	≈	×	.35 (.16)	.30 (.13)	>	✓	.35 (.16)	.30 (.13)	>	✓
Vs.	coop→indiv	.39 (.23)	.33 (.24)	>	*F2.1*	.33 (.14)	.32 (.14)	≈	×	.35 (.15)	.31 (.14)	>	✓	.33 (.14)	.29 (.13)	>	✓
Indiv	indiv→coop	.38 (.26)	.33 (.24)	>	*F2.1*	.33 (.15)	.33 (.13)	≈	×	.35 (.16)	.30 (.13)	>	✓	.38 (.18)	.31 (.13)	>	✓
		Cooperative	*Control*			*Cooperative*	*Control*			*Cooperative*	*Control*			*Cooperative*	*Control*		
Coop	both	.30 (.22)	.26 (.22)	>	*F2.2a*	.33 (.14)	.34 (.14)	≈	×	.36 (.16)	.31 (.12)	>	✓	.35 (.16)	.30 (.13)	>	✓
Vs.	coop→ctrl	.30 (.20)	.31 (.23)	≈	*F2.2b*	.32 (.14)	.33 (.14)	≈	✓	.36 (.15)	.30 (.12)	>	×	.34 (.14)	.33 (.15)	≈	✓
Ctrl	ctrl→coop	.31 (.27)	.13 (.15)	>	*F2.2b*	.34 (.15)	.35 (.15)	≈	×	.37 (.16)	.31 (.12)	>	✓	.37 (.17)	.27 (.11)	>	✓
		Individual.	*Control*			*Individual.*	*Control*			*Individual.*	*Control*			*Individual.*	*Control*		
Indiv	both	.33 (.15)	.30 (.19)	≈	*F2.3*	.33 (.14)	.33 (.14)	≈	✓	.32 (.13)	.31 (.13)	≈	✓	.30 (.12)	.28 (.11)	≈	✓
Vs.	indiv→ctrl	.35 (.15)	.31 (.19)	≈	*F2.3*	.32 (.14)	.34 (.14)	≈	✓	.31 (.13)	.30 (.13)	≈	✓	.30 (.13)	.29 (.11)	≈	✓
Ctrl	ctrl→indiv	.31 (.15)	.29 (.20)	≈	*F2.3*	.33 (.15)	.33 (.14)	≈	✓	.32 (.13)	.32 (.14)	≈	✓	.31 (.11)	.28 (.11)	>	×
Study 3		*Cooperative*	*Competitive*			*Cooperative*	*Competitive*			*Cooperative*	*Competitive*			*Cooperative*	*Competitive*		
Coop	both	.41 (.23)	.39 (.21)	≈	*F3.1a*	.34 (.15)	.34 (.15)	≈	✓	.36 (.15)	.33 (.13)	>	✓	.39 (.16)	.35 (.15)	>	✓
Vs.	coop→comp	.37 (.18)	.49 (.19)	<	*F3.1b*	.35 (.15)	.33 (.14)	≈	×	.36 (.15)	.32 (.14)	>	×	.33 (.13)	.38 (.16)	<	✓
Ctrl	comp→coop	.44 (.25)	.32 (.20)	>	*F3.1c*	.34 (.14)	.34 (.15)	≈	×	.37 (.16)	.33 (.13)	>	✓	.46 (.17)	.31 (.12)	>	✓
		Cooperative	*Control*			*Cooperative*	*Control*			*Cooperative*	*Control*			*Cooperative*	*Control*		
Coop	both	.34 (.17)	.24 (.14)	>	*F3.2a*	.34 (.14)	.34 (.15)	≈	×	.36 (.14)	.31 (.13)	>	✓	.38 (.16)	.31 (.14)	>	✓
Vs.	coop→ctrl	.24 (.09)	.21 (.12)	≈	*F3.2b*	.34 (.14)	.34 (.14)	≈	✓	.35 (.14)	.32 (.13)	>	×	.32 (.12)	.34 (.16)	≈	✓
Ctrl	ctrl→coop	.39 (.19)	.26 (.15)	>	*F3.2b*	.33 (.14)	.33 (.15)	≈	×	.36 (.14)	.31 (.13)	>	✓	.44 (.17)	.29 (.12)	>	✓
		Competitive	*Control*			*Competitive*	*Control*			*Competitive*	*Control*			*Competitive*	*Control*		
Comp	both	.23 (.11)	.23 (.17)	≈	*F3.3*	.35 (.15)	.34 (.14)	≈	✓	.33 (.13)	.31 (.13)	≈	✓	.29 (.11)	.29 (.11)	≈	✓
Vs.	comp→ctrl	.22 (.10)	.25 (.18)	≈	*F3.3*	.35 (.15)	.35 (.14)	≈	✓	.33 (.12)	.31 (.12)	≈	✓	.30 (.11)	.29 (.10)	≈	✓
Ctrl	ctrl→comp	.25 (.13)	.20 (.16)	≈	*F3.3*	.35 (.14)	.34 (.14)	≈	✓	.32 (.13)	.32 (.13)	≈	✓	.27 (.10)	.29 (.11)	≈	✓

References

1. Jennings, N. R., Sycara, K., Wooldridge, M.: A Roadmap of Agent Research and Development. Autonomous Agents and Multi-Agent Systems 1, 275--306 (1998)
2. Kraus, S.: Negotiation and Cooperation in Multi-Agent Environments. Artificial Intelligence 94(1-2), 79--98 (1997)
3. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. Science 211, 453--458 (1981)
4. Loewenstein, G., Lerner, J.: The role of affect in decision making. In: Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (eds.), Handbook of Affective Sciences, pp.619--642. Oxford University Press (2003)
5. Kollock, P.: Social Dilemmas: The Anatomy of Cooperation. Annual Review of Sociology, 24, 183--214 (1998)
6. McClintock, C.G., Liebrand, W.B.G.: Role of interdependence structure, individual value orientation, and another's strategy in social decision making: a transformational analysis. J. Pers. Soc. Psychol. 55(3), 396--409 (1988)
7. Kramer, R.M., Brewer, M.B.: Social group identity and the emergence of cooperation in resource conservation dilemmas. In: Wilke, H.A.M., Messick, D.M., Rutte, C. (eds.), Experimental Social Dilemmas, pp. 205--34. Verlag Peter Lang, Frankfurt (1986)
8. Axelrod, R.: The Evolution of Cooperation. Basic Books, New York (1984)
9. Yamagishi T.: The provision of a sanctioning system as a public good. J. Pers. Soc. Psychol. 51, 110--116 (1986)
10. Jerdee, T.H., Rosen, B.: Effects of opportunity to communicate and visibility of individual decisions on behavior in the common interest. J. Appl. Soc. Psychol. 59, 712--16 (1974)
11. Boone, R., Buck, R.: Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. J. of Nonverbal Behav., 27, 163--182 (2003)
12. Frank, R. H.: Passions within reason: The strategic role of the emotions. Norton, NY (1988)
13. Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T. Bonnet, K.: Emotional expressivity as a signal of cooperation. Evolution and Human Behavior, 31, 87--94 (2010)
14. de Melo, C., Zheng, L., Gratch, J.: Expression of Moral Emotions in Cooperating Agents. In: Proceedings of Intelligent Virtual Agents 2009, pp. 301--307 (2009)
15. de Melo, C., Carnevale, P., Gratch, J.: The influence of Emotions in Embodied Agents on Human Decision-Making. In: Proc. of Intelligent Virtual Agents 2010, pp. 357--370 (2010)
16. de Melo, C., Carnevale, P., Gratch, J.: The Impact of Emotion Displays in Embodied Agents on Emergence of Cooperation with People. Submitted to J. Presence (submitted)
17. de Melo, C., Carnevale, P., Gratch, J.: Reverse Appraisal: Inferring from Emotion Displays who is the Cooperator and the Competitor in a Social Dilemma. In Proc. of 33rd Annual Meeting of the Cognitive Science Society, pp. 396–401 (2011)
18. Gratch, J., Rickel, J., Andre, E., Badler, N., Cassell, J., Petajan, E. Creating Interactive Virtual Humans: Some Assembly Required. IEEE Intelligent Systems 17(4), 54--63 (2002)
19. Reeves, B., Nass, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. University of Chicago Press (1996)
20. Blascovich, J.: Social influence within immersive virtual environments. In: Schroeder, R. (ed.), The social life of avatars: Presence and interaction in shared virtual environments, pp. 127--145. Springer-Verlag, London (2002)
21. Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., Bailenson, J. N. Immersive virtual environment technology as a methodological tool for social psychology. Psychological Inquiry 13, 103--124 (2002)
22. Hilty, J., Carnevale, P. Black-Hat/White-Hat Strategy in Bilateral Negotiation. Organizational Behavior and Human Decision Processes, 55, 444--469 (1993)