Emotional Signaling in a Social Dilemma: an Automatic Analysis

Giota Stratou, Rens Hoegen, Gale Lucas and Jonathan Gratch
University of Southern California, Institute for Creative Technologies
Los Angeles, California
Email: {stratou, rhoegen, lucas, gratch }@ict.usc.edu

Abstract—**Emotional signaling plays an important role in negotiations and other social decision-making tasks as it can signal intention and shape joint decisions. Specifically it has been shown to influence cooperation or competition. This has been shown in previous studies for scripted interactions that control emotion signaling and rely on manual coding of affect. In this work we examine face-to-face interactions in an iterative social dilemma task (prisoner's dilemma) via an automatic framework for facial expression analysis. We explore if automatic analysis of emotion can give insight into the social function of emotion in face-to-face interactions. Our analysis suggests that positive and negative displays of emotion are associated with more prosocial and proself game acts respectively. Moreover signaling cooperative intentions to the opponent via positivity can leave participants more open to exploitation, whereas signaling a more tough stance via negativity seems to discourage exploitation. However, the benefit of negative affect is short-term and both players do worse over time if they show negative emotions.**

Keywords—*emotional signaling; iterated prisoner's dilemma; automatic analysis; social signaling*

I. INTRODUCTION

Imagine you are a North Vietnamese diplomat heading to Paris to negotiate a cease-fire with the United States. Your American counterpart calls you to say "for God's sake, you know Nixon is obsessed about communism. We can't restrain him when he's angry-and he has his hand on the nuclear button" ([1, p. 83]). Would this affect your behavior? Or imagine a more mundane example. You are playing poker with a group of friends. After drawing a card, one of the other players flashes a broad smile. Would this affect your behavior? In each of these cases, another person's expressed emotion affects our impression of the motives behind his or her actions, shaping our own emotions and decisions, and the other party's subsequent reactions to us. Mutual understanding or conflict between individuals, groups, and even nations can depend on such emotional signaling. Affective computing techniques hold the potential to recognize, understand and even shape these emotional processes.

In this paper, we examine emotional signaling in face-to-face interactions and determine if automatic analysis of emotional signals can give insights into how parties reach social decisions. Specifically, we consider how displays of emotion shape joint decisions in the prisoner's dilemma, a standard laboratory task for studying trust and cooperation. Previous research in psychology has highlighted the important role of facial expressions in shaping such joint decisions. We

seek to extend this research in two important ways. First, most prior research has involved the use of confederates (i.e., people interact with a human or computer program that exhibits a scripted sequence of emotional displays), whereas here we seek to replicate these findings in the context of unscripted dyadic interactions. Second, most psychological research uses laborious manual annotations of emotional displays, whereas here we employ computer vision techniques to classify emotional signals.

Prior psychological and human-computer research using confederates has revealed a number of robust findings on how emotions shape social decisions. For example, de Melo and colleagues had participants play the iterated prisoner's dilemma with an expressive "virtual confederate" [2]. They showed that display of guilt after exploiting the participant elicited more pro-social responses from participants than an agent that smiled. In the context of an ultimatum game, Schug and colleagues manually annotated the expressions of participants and found that emotional expressivity was a reliable signal of cooperativeness [3]. In yet another example, Krumhuber and colleagues used videotaped confederates and virtual humans to demonstrate how the temporal dynamics of a smile can shape decisions in a trust game [4].

One of the more robust findings on emotional signaling stems from the research of Van Kleef and colleagues, and our present research builds on these seminal findings [5]. In this work, they examined the social influence of positive and negative expressions in a competitive setting. Specifically, the study used a computer-mediated multi-issue negotiation scenario, where participants faced an opponent that expressed anger, happiness or nothing (control). Participants were carefully led to believe they are negotiating with another participant, through a computer, but in fact they were matched with a computer program that plays a scripted strategy. The results showed that participants conceded more when facing an angry opponent than the control and participants conceded less when matched with the happy opponent than the control. Van Kleef argues that participants are using emotion to infer their opponent's goals and intentions. When faced with an angry opponent, they estimate the opponent will be uncompromising (i.e., is tough) and, to avoid a costly impasse, they make large concessions. When faced with a happy opponent, they infer the opponent is compromising (i.e., soft) and, thus, demand high concessions from them. *In other words, a happy opponent*

IEEE
computer
society

		Opponent Picks	
		Cooperate (*C*)	Defect (*D*)
Participant Picks	Cooperate (*C*)	5,5	0,10
	Defect (*D*)	10,0	1,1

invites exploitation whereas an angry opponent avoids being exploited.

This, perhaps paradoxical finding that happy displays lead to exploitation, at least in competitive contexts, has been replicated in several studies and is one of the more robust findings on the social consequences of emotion displays. It has even been replicated in the context of human-computer interaction. For example, de Melo showed that people exploited a smiling virtual human, but were exploited by a virtual human showing anger, even when they knew they were playing against a computer program [2]. Nonetheless, for all studies that we are aware of, this phenomenon has only been shown in artificial interactions, where the expressions of one party are completely scripted.

Here we examine this finding in natural dyadic interactions between two participants in the iterated prisoner's dilemma task. We explore if automatic analysis of emotion can give insights into the social function of emotion in face-to-face interactions. In the next section, Section II, we will review the prisoner's dilemma task and our hypotheses based on the nature of the task and previous literature. In Section III, we give an overview of our data and experimental setup. Section IV describes the behaviors we encoded automatically, following up with results and discussion in Sections V and VI respectively. Finally we close with conclusions in Section VII.

II. EMOTIONS IN THE PRISONER'S DILEMMA

The prisoner's dilemma [6] is a two-player task where the payoffs of each player depend on the simultaneous choice of both players. Each player must decide between two possible options. One option is considered the cooperative choice as it leads to the highest payout if both players select this option, however this choice also places a player at risk for exploitation. If one player chooses the cooperative choice and the other choses the alternative non-cooperative choice (conventionally referred to as the *Defect* choice), the defector gets a large payoff and the cooperator a small one. If both players defect, they receive a small payout. The payoff matrix used in our experiment is shown in Table I. The task represents a dilemma because the rational (i.e., utility-maximizing) choice for both players is to defect, which again results in a payout (mutual defection) that is worse than mutual cooperation.

The iterated prisoner's dilemma is a simple extension of the standard game where the same players play multiple rounds with each other. This creates interesting and complex dynamics. For example, if a player is exploited in one round, they can try to punish the offender in the following round. Figure 1 illustrates the various states and actions that can arise
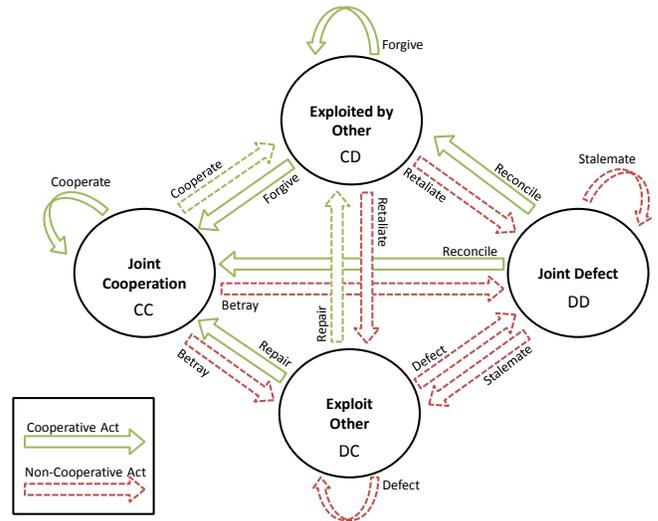

Fig. 1. A network of game states and all possible transitions in the iterative prisoner's dilemma task. The transitions, or game acts, have been labeled from a player-centric point of view and classified in cooperative acts (choose *Cooperate* in a certain round) and non-cooperative acts (choose *Defect* in a certain round)

in the iterated prisoner's dilemma. If both players choose the cooperative choice they enter the state of *Joint Cooperation*. From this state, they can continue to cooperate, one player can betray the other, or both could simultaneously transition to *Joint Defection*. Interestingly, allowing multiple rounds does not change the rational strategy. As long as players know there is a finite number of rounds, the rational choice is to defect, which again results in an outcome that is worse than mutual cooperation.

Here we are interested in how expressions of emotion will impact how players solve this iterated dilemma. Following the findings of Van Kleef [5], positive emotional displays should signal that the player will be cooperative, and should invite exploitation by the other player. In contrast, negative emotional displays should signal that the player will be tough and will avoid exploitation. Based on this, we develop the following hypotheses:

H1a: *Positive emotions on the part of one player will invite exploitation by the other player*

H1b: *Negative emotions will discourage exploitation by the other player*

Following Van Kleef, we conjecture this occurs because positive emotions signal prosocial actions on the part of a player. Specifically, we hypothesize positive players will show more cooperative acts in general:

H2a: *Positive emotions signal cooperative actions, meaning the probability that a player will make the cooperative choice, regardless of the choice of the opponent, is positively correlated with positive displays of emotion.*

H2b: *Negative emotions on the part of one player are associated with non-cooperative actions, meaning that the probability that a player makes the non-cooperative choice, regardless of the choice of the opponent, is positively correlated with negative emotion.*
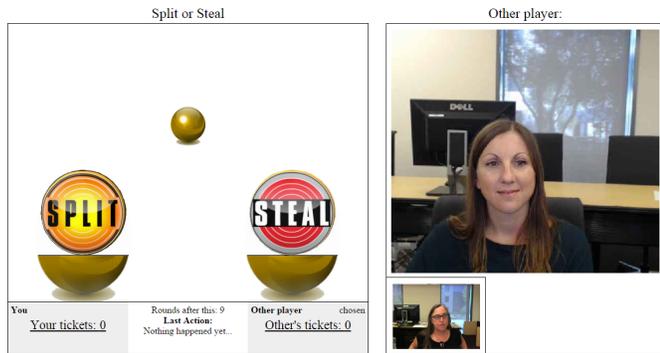
Fig. 2. A view of the player's screen during the task. On the left side one can see the game interface, where the player can make their decision for the current round and is given info about status (rounds left, number of tickets won by them and by opponent). On the right side, the player can see the opponent's camera in the big window and their camera on the small window on the bottom (similar to a chat-like interface)

These hypotheses concern individual decisions, however, the outcome of the overall game is more complex and emerges from the interactions between players. For example, a player that is exploited may subsequently retaliate. Although we have no specific hypotheses, we also investigate how emotional expressions relate to players' scores over the multiple rounds.

Finally, previous research [7] shows that people have different dispositional tendencies towards cooperation and competition, thus we measure these dispositions before the game using the social value orientation scale and examine if the patterns of expression differ depending on whether participants are predisposed to be prosocial or proself.

III. Experimental Setup

In this section we will describe the data and population statistics.

A. Configuration

We elicited facial displays of emotion by having pairs of participants play a computer-mediated version of the iterated prisoner's dilemma game shown in Figure 2. The game was patterned after a version of the prisoner's dilemma used in the British TV show *Golden Balls*. In this version, the cooperative choice is called *Split* and the defect choice is called *Steal*. We adopted this design after several pilot studies as it makes the structure of the task more intuitively obvious and the game-like framing seems to elicit more nonverbal communication between participants than other more conventional framings.

The participants played in pairs in a laboratory setting, seated in front of a computer with a webcam. Before starting the task, participants were given some time to introduce themselves to one another to break the ice. Participants could pick their game acts by clicking on the screen and they could see each other by camera in a chat-like interface next to the game. Groups of participants were paired up randomly and seated in separate cubicles. They were instructed not to talk and could only see or interact with each other via the interface. The participants were instructed to play 10 rounds of the game with the payoffs shown in Table I. In this case,

the payoff was translated into lottery tickets that would be entered into a main lottery with a $100 prize. Participants were given 10 lottery tickets at the beginning of the game and motivated to win as many lottery tickets as possible throughout the game by making a choice of splitting with their opponent or stealing from them at every round (following the *Cooperate* and *Defect* payoffs from Table I). All participants were paid a basic participation fee of $30 for playing the game.

The game acts of the participants were logged in a database along with the timestamps. Using this data we could infer when and what decisions were made, when the reveal of the joint result happened and when each round began and ended.

B. Population

In total the analyzed data includes 186 participants from the Los Angeles metropolitan area (93 pairs), who were recruited using Craigslist and met some basic requirements (age, language, adequate eyesight). Before the game, the participants completed several questionnaires including a 5 point scale MACH-IV [8] measuring Machiavellian personality, a 7 point scale Emotion Regulation Questionnaire (ERQ) [9] measuring tendencies to suppress and reappraise emotions, a 7 point scale Berkeley Expressivity Questionnaire (BEQ) [10] measuring tendency to express emotions and the Triple Dominance Measure of Social Value Orientation questions [11] which can be used to classify individuals as prosocial (i.e., concerned with the needs of others) or proself (i.e., self-interested). They were also asked to fill in a post-game questionnaire about their feelings and impressions (self and of opponent) based on the total interaction.

Based on the Social Value Orientation (SVO) instrument we were able to classify participants into *prosocial* (S) and *proself* (I). In our analyzed data we had 106 (out of 186) participants in the prosocial class.

IV. Automatic Behavior Extraction

To measure emotional signaling related to game behaviors and outcome we proceed to encode both from our data. For the game actions and outcome we have exact, quantifiable measures taken from the game logs. For the emotional signaling we will examine emotional displays as measured from the participant videos, by automatic methods for expression recognition.

A. Game Behavior

In Figure 1 we show the game states and all possible game acts for the iterative prisoner's dilemma task. We encoded both static measures over the whole game, such as the number of times that a state happened (Ex. count of Joint Cooperation, or *CC*, states) and also player game acts, (Ex. *Betrayals*, defined as the number of times that a player is choosing to leave a joint cooperation state by defecting). The naming of transitions follows the Matsumoto [12] terminology for the player's game acts. The game acts were measured as probabilities a player will act in that manner. For example, *BetrayChance* is the number of times a player chose to *Betray*, divided by the

Y	X	RELATIONSHIP			
		All	Prosocial(S)	Proself(I)	
GetBetrayedChance	POS_self	↗**	↗**		H1a
CooperativeActChance	POS_self		↗ +		H2a
RepairChance	POS_self		↗ *		H2a
ReconcileChance	POS_self		↗ +		H2a
NonCooperativeActChance	POS_self		↘ +		H2a
DefectChance	POS_self		↘ *		H2a
StalemateChance	POS_self		↘ +		H2a
CD counts	POS_self	↗ +	↗ +		H1a
DD counts	POS_self		↘ *		
PlayerScoreDiff	POS_self	↘ +	↘ *		
PlayerScoreRatio	POS_self		↘ *		
OpponentScore	POS_self	↗ *	↗ **		
DyadScore	POS_self		↗ *		
GetBetrayedChance	NEG_self		↘ +		H1b
NonCooperativeActChance	NEG_self	↗ *			H2b
BetrayChance	NEG_self	↗ *			H2b
RetaliateChance	NEG_self	↗ *			H2b
CooperativeActChance	NEG_self	↘ *			H2b
ForgiveChance	NEG_self	↘ *			H2b
CooperateChance	NEG_self	↘ *			H2b
DD counts	NEG_self	↗ *	↗ *		H2b
PlayerScoreRatio	NEG_self		↗ +		
OpponentScore	NEG_self	↘ *	↘ +		
DyadScore	NEG_self	↘ *	↘ *		
GetBetrayedChance	POS_opponent			↗ *	

amount of times that this opportunity was given (that the player was in a *JointCooperation* state). On top of these measures, we added the only game act that is not purely chosen by the player *GetBetrayedChance* (which is a receiving act) because it directly relates to our hypotheses. In addition to the behaviors, we also measure the game outcome for both the player and the dyad by adding the following measures: {*PlayerScore*, *OpponentScore*, *DyadScore*, *PlayerScoreDiff*, *PlayerScoreRato*}. *DyadScore* is the score of both players together, *PlayerScoreDiff* is the relative difference of the player's score to the opponent's score and *PlayerScoreRatio* is the ratio of the player's score over the dyad score.

B. Automatic Extraction of Emotional Displays

Participant videos from the webcams were automatically analyzed using FACET [13] facial expression recognition software.[1] FACET features include intensities for the basic emotion labels including "NEUTRAL" as well as for the basic sentiment classes: "POSITIVE" and "NEGATIVE".[2] For this work we will focus only on these three main labels because

it simplifies our first analysis.[3]

Logging of the game events allowed for automatic event-based behavior encoding as well as automatic segregation of the signals on the game period from the overall recording. We encoded the sentiment labels in three different scopes of the game: Full Round, Play (the time between the start of the round till the player makes a decision) and Reveal (around the time the joint decision is revealed), then average over all rounds for an overall game profile. The reason why we encoded in those separate areas is to separate potential effects on times of the game where they can have different significance, based on a scheme proposed by Matsumoto for analyzing dyadic affect[12]. When summarizing behaviors over all 10 rounds, however, these measures on different game scopes are highly correlated (with correlation coefficients >0.9), so in our analysis we will present results only for the full round averages. For convenience, from now on we will refer to

---

[1] http://www.emotient.com/products/#FACETVision

[2] These labels have been trained on thousands of images, annotated by experts to be conveying positive or negative emotion, allowing overlap

[3] We also report the average per frame correlation of the automatic measures over all participants on our data: corr(POS, NEG) = -0.53, corr(POS, NEU) = -0.34 and corr(NEG, NEU) = -0.11. Also, we found NEGATIVE correlated positively with ANGER, DISGUST, SADNESS, AU4 and negatively with JOY. POSITIVE correlated positively with JOY, AU12, AU6 and negatively with ANGER, SADNESS, CONFUSION, showing that they follow intuitive relationships

these as: {*POS_self*, *NEG_self*, *NEU_self*, *POS_opponent*, *NEG_opponent*, *NEU_opponent*} for the participant and the opponent respectively.

Videos and segments with high rate of missing frames were automatically discarded from the analysis.

V. ANALYSIS

We look at the relationship between overall game behaviors and emotional displays. Our aim is that such analysis will give insights into the social process that happens during an iterative game and especially with observations related to our hypotheses.

A. Statistical Analysis

Specifically, we will test if emotional display exchanges can predict players' choices in the game. For that purpose, we will use a simple linear regression model [4]: Y∼1+X, to model the relationship of each of our emotional display variables X with a game behavior or outcome measure Y. We summarize the results in Table II using the ↗ and ↘ notation to show increasing and decreasing relationships respectively. For significance levels the notation is as follows: + p<=0.1, * p<=0.05, ** p<=0.01. For example, the first line of the table shows us that the more positive a player's emotional display was overall, the more the chance of getting exploited, and this is significant at the 0.01 level for *all* and *prosocial* population.

B. Results

All results are summarized in Table II. Here we discuss some of the significant findings for the population as a whole.

On H1: We find partial support for our hypothesis that emotion displays impact the chance of getting exploited. The chance of getting exploited is significantly increasing (coeff=0.275, p=0.0015) with increasing positive displays (*H1a*). Also, the number of times an exploitation state happens to a player (*CD*) is correlated positively (coeff=0.902, p=0.0937) with the amount of positivity they display (this result was also shown symmetrically from the opponent's side with the relationship *DC∼1+POS_opponent*). However, we did not find support that negative displays inoculate players from exploitation (*H1b*).

On H2: We find good support for the hypothesis that emotion displays predict social act. Specifically, hypothesis *H2b* that negative displays are associated with non-cooperative gaming behavior was shown both directly, with the relationship *NonCooperativeActChance∼1+NEG_self* (coeff=0.196, p=0.0520), and with specific instances for *BetrayChance*, *RetaliateChance* (both increase with more negative displays, (coeff=0.270, p=0.0371) and (coeff=0.301, p=0.0357) respectively). We also find that the number of *Joint Defection* (*DD*) states increases (coeff=1.574, p=0.0276) with negativity. *H2a* though, was not confirmed directly for the whole population since we did not observe any significant effect of *POS_self* on any cooperative act.

[4]http://www.mathworks.com/help/stats/linearmodel-class.html
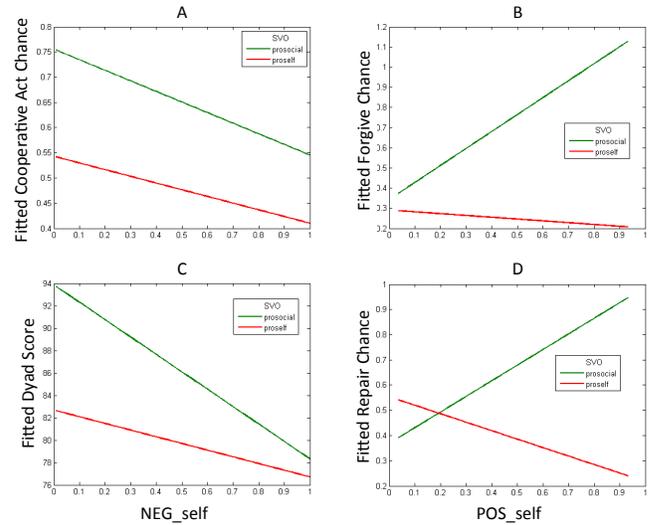
Fig. 3. Some examples of SVO interaction with emotional display in the prediction of game behaviors and outcome. In cases A and C, these models reported no significant interaction with SVO (behaviors are consistent among SVO categories of prosocial and proself). In contrast, models B and D reported a significant interaction between positive display and SVO in the prediction of certain cooperative acts.

On game outcome: The observation that positivity invites exploitation does not necessarily mean that positive players perform poorly as the game persists over multiple rounds. Indeed, we found several interesting associations between emotional displays and the overall score for both individuals and the dyad. On the one hand, we find evidence suggesting that positivity opens the door to exploitation, whereas negativity signals toughness and thus discourages exploitation. Specifically, we find that if a player shows positive emotion, their opponent scores better (coeff=9.673, p=0.0507). Similarly, if a player shows negative emotions, their opponent scores worse (coeff=-8.936, p=0.0376). However, even though they boost the relative (individual) gain, in the long run negative displays seem to hurt the joint payoff: as a dyad, both players do worse if one of the players is more negative (coeff=-12.595, p=0.0276). This implies that the short-term benefits of negative affect may not translate into long-term gains. This may be because of the iterative nature of the dilemma, where negative displays and non-cooperative play may backfire.

C. Interaction with Social Value Orientation

To gain further insight into player behavior, we examine if these results differ depending on the player's SVO disposition (prosocial or proself). We extend the previous analysis by separating the populations based on SVO and by testing the relationships separately in the two subpopulations prosocial (S) and proself (I). We present the results of the extended analysis for all three different populations: *All*, *Prosocial* participants only (S), *Proself* participants only (I) in Table II. This extended analysis reveals that the main effects on the whole population are driven mainly by the prosocial group of participants. Looking at the subpopulation of prosocial

participants we find support for most of our original hypotheses. Specifically, hypothesis *H2a* that cooperative game acts are associated with displays of positivity holds for that subpopulation (coeff=0.292, p=0.0527).

As an additional test of the effect of SVO on the relationship of displays with game behavior, we test the linear regression models on the whole population and we introduce the SVO as a parameter to the regression model. These tests yield consistent findings with what was presented in Table II and furthermore allow us to look closer at the interaction of SVO with the emotion displays when predicting aspects of game behavior. As an example we show some of the model effects with SVO as a parameter in Figure 3. Negative expression has a main effect on the whole population and there was no significant interaction with SVO, whereas positive interacts significantly with SVO to predict some of the prosocial acts (for example, $RepairChance\sim1+SVO*POS\_self$ (coeff=-0.956, p=0.0074)).

VI. Discussion

This is a first step towards automatic analysis of emotion in a social dilemma task. Via automatic expression recognition methods we were able to capture participants' displays of positivity, negativity and neutrality from videos of real face-to-face interactions during an iterative prisoner's dilemma task. Our main observations support previous findings [5], [2] from both human interactions and human-computer interactions, that displays of positivity in a competitive scenario can signal cooperative intention and thus open the door to exploitation, whereas negative displays can signal a tough stance and thus discourage exploitation.

We also presented evidence on the overall game outcome showing that even though negativity boosts the relative individual score in the short-term, this may not translate into long-term benefits. This is also supported by literature: negative affect in negotiations decreases joint gains [14] and communicating toughness via anger affects mostly recipients that have poor alternatives in a negotiation [15]. In our case, iterating over multiple rounds gives participants more alternatives, so the dynamics may change.

The SVO interaction analysis also revealed some interesting results. It seems that negative displays have a consistent effect on game behaviors independently of the participant being prosocial or proself, but it is not the same case with positive displays which seem to be driven mostly by prosocial participants.

Our current analysis brings interesting insights into the relationship between behaviors and emotional displays in the iterative prisoner's dilemma task, but it does so on an overall level that makes it difficult to decouple causality. Specifically, we showed evidence that emotional signals are associated with game behaviors the way we would expect from literature but since this is an iterative task could they also be a result of the previous game outcome? Is positivity causing more prosocial acts? or is it a result of a more cooperative and thus satisfying game? These interesting questions are best addressed by studies that experimentally manipulate emotion in a systematic way. Virtual confederates (i.e., digital characters that appear to be human but can precisely manipulate their nonverbal behavior) are one promising approach to unpack this causality (see [16] and [17]). We plan to use this approach in our future work.

Another point for discussion is the use of the overall sentiment labels *POSITIVE, NEGATIVE* instead of individual expression labels like *Joy*, *Contempt*,*Anger* etc. or individual facial action units. We would argue that Positive is highly correlated with Joy (and perhaps imposes fewer restrictions on interpretation whether we observe real or social smiles) on an overall level. In the case of Negative, we agree that there is more variation on what expression can be classified as negative (usually it includes anger, contempt, disgust, and sometimes sadness and fear) and wider valence range so in that case the differentiation could perhaps shed more light into future analysis.

VII. Conclusion

We presented evidence from real face-to-face interactions that game behaviors in an iterative social dilemma task are associated with emotional signal exchanges. We validated behavioral theories from literature in more natural interactions and via automatic analysis (avoiding tedious manual annotations). We would argue that the interactions here are more challenging in that perspective, since the emotional signals are not controlled by experimental framing and are thus noisier. We also gave some insights into new directions factoring in SVO. Besides providing a new framework for further analysis, this work has direct implications for affective computing, since it paves the way towards automatic systems that can automatically recognize the emotion signals from a human interlocutor and assign an intent or meaning in the context of a social dilemma or negotiation task. The proposed direction of this work is to replicate such emotional exchanges in a human-machine interaction [5], which would be particularly useful for negotiation training systems (e.g. [18], [19]) that could both perceive the players' intention from their nonverbal behaviors (some promising work has been done on that [20]) and show their intentions as well (e.g. by emotion displays on a virtual human agent) in a manner similar to human-to-human interactions.

[5]Our techniques and recognition methods are automatic and real-time implementable facilitating the extension of those models of learned behavior to a real-time automatic system with a virtual agent

IEEE
computer
society

REFERENCES

[1] H. R. Haldeman and J. DiMona, *The ends of power*. Times Books New York, 1978.

[2] C. M. de Melo, P. Carnevale, and J. Gratch, "The effect of expression of anger and happiness in computer agents on negotiations with humans," in *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3*, ser. AAMAS '11. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2011, pp. 937–944. [Online]. Available: http://dl.acm.org/citation.cfm?id=2034396.2034402

[3] J. Schug, D. Matsumoto, Y. Horitaa, T. Yamagishi, and K. Bonnet, "Emotional expressivity as a signal of cooperation," *Evolution and Human Behavior*, vol. 31, pp. 87–94, 2010.

[4] E. Krumhuber, A. Manstead, D. Cosker, D. Marshall, P. L. Rosin, and A. Kappas, "Facial dynamics as indicators of trustworthiness and cooperative behavior," *Emotion*, vol. 7, no. 4, pp. 730–735, 2007.

[5] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead, "The interpersonal effects of anger and happiness in negotiations." *Journal of personality and social psychology*, vol. 86, no. 1, p. 57, 2004.

[6] W. Poundstone, *Prisoner's Dilemma*, 1st ed. New York, NY, USA: Doubleday, 1993.

[7] P. A. Van Lange, "The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation." *Journal of Personality and Social Psychology*, vol. 77, no. 2, p. 337, 1999.

[8] R. Christie and F. L. Geis, *Studies in machiavellianism*. Academic Press, 2013.

[9] J. J. Gross and O. P. John, "Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being." *Journal of personality and social psychology*, vol. 85, no. 2, p. 348, 2003.

[10] J. J. Gross and O. P. John, "Revealing feelings: facets of emotional expressivity in self-reports, peer ratings, and behavior." *Journal of personality and social psychology*, vol. 72, no. 2, p. 435, 1997.

[11] P. A. Van Lange, E. De Bruin, W. Otten, and J. A. Joireman, "Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence." *Journal of personality and social psychology*, vol. 73, no. 4, p. 733, 1997.

[12] D. Matsumoto, "Facial affect reciprocity in dyadic interactions," DTIC Document, Tech. Rep., 2012.

[13] G. Littlewort, J. Whitehill, T. Wu, I. Fasel, M. Frank, J. Movellan, and M. Bartlett, "The computer expression recognition toolbox (cert)," in *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, 2011, pp. 298–305.

[14] K. G. Allred, J. S. Mallozzi, F. Matsui, and C. P. Raia, "The influence of anger and compassion on negotiation performance," *Organizational Behavior and Human Decision Processes*, vol. 70, no. 3, pp. 175 – 187, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0749597897927059

[15] M. Sinaceur and L. Z. Tiedens, "Get mad and get more than even: When and why anger expression is effective in negotiations," *Journal of Experimental Social Psychology*, vol. 42, no. 3, pp. 314 – 322, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022103105000703

[16] C. M. de Melo, P. J. Carnevale, and J. Gratch, "Using virtual confederates to research intergroup bias and conflict," *Academy of Management Proceedings*, vol. 2014, no. 1, p. 11226, 2014.

[17] J. Blascovich, J. Loomis, A. Beall, K. Swinth, C. Hoyt, and J. N. Bailenson, "Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry*, vol. 13, pp. 103–124, 2002.

[18] J. Broekens, M. Harbers, W.-P. Brinkman, C. M. Jonker, K. Van den Bosch, and J.-J. Meyer, "Virtual reality negotiation training increases negotiation knowledge and skill," in *Intelligent Virtual Agents*. Springer, 2012, pp. 218–230.

[19] J. Kim, R. W. Hill, P. Durlach, H. C. Lane, E. Forbell, M. Core, S. C. Marsella, D. V. Pynadath, and J. Hart, "BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context," *International Journal of Artificial Intelligence in Education*, vol. 19, no. Issue on Ill-Defined Domains, pp. 289–308, 2009. [Online]. Available: http://ict.usc.edu/pubs/BiLAT-%20A%20Game-Based%20Environment%20for%20Practicing%20Negotiation%20in%20a%20Cultural%20Context.pdf

[20] S. Park, S. Scherer, J. Gratch, P. Carnevale, and L.-P. Morency, "Mutual behaviors during dyadic negotiation: Automatic prediction of respondent reactions," in *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, 2013, pp. 423–428.

IEEE
computer
society