The Effect of Agency on the Impact of Emotion Expressions on People's Decision Making

Celso M. de Melo
Marshall School of Business
University of Southern California
Los Angeles, USA
demelo@usc.edu

Jonathan Gratch
Institute for Creative Technologies
University of Southern California
Los Angeles, USA
gratch@ict.usc.edu

Peter J. Carnevale
Marshall School of Business
University of Southern California
Los Angeles, USA
peter.carnevale@marshall.usc.edu

***Abstract*— Recent research in neuroeconomics reveals that people show different behavior and lower activation of brain regions associated with mentalizing (i.e., the inference of other's mental states) when engaged in decision making tasks with a computer, when compared to a human. These findings are important for affective computing because they suggest people's decision making might be influenced differently according to whether they believe the emotional expressions shown by a computer are being generated by a computer algorithm or a human. To test this, we had people engage in a social dilemma (Experiment 1) or a negotiation (Experiment 2) with virtual humans that were either agents (i.e., controlled by computers) or avatars (i.e., controlled by humans). The results show a clear agency effect: in Experiment 1, people cooperated more with virtual humans that showed facial cooperative displays (e.g., joy after mutual cooperation) rather than competitive displays (e.g., joy when the participant was exploited) but, the effect was only significant with avatars; in Experiment 2, people conceded more to an angry than a neutral virtual human but, once again, the effect was only significant with avatars.**

Keywords—Agency; Emotion Expression; Decision Making

I. INTRODUCTION

There has been growing interest in the behavioral sciences on the effects of emotion expressions on others' decision making [1]. Acknowledging that emotions help regulate social interaction and serve important social functions such as communicating one's beliefs, desires and intentions to others [2], [3], [4], the social effects of emotions displays on people's decision making have now been reported in negotiation (e.g., [5], [6]), trust games (e.g., [7]), ultimatum games (e.g.,[8]), public goods dilemma (e.g., [9]), dispute resolution [10] and daily life [11]. From a computational perspective, it is important we understand whether emotion expressions in computers can also influence people's decisions and researchers have already begun exploring this question. Indeed, in recent studies, we have shown that emotion expressions in virtual humans can impact people's decision to cooperate in a social dilemma [12] and people's concession making in negotiation [13]. It would thus seem that the social effects of emotion displays we see in human-human interaction straightforwardly carry to human-agent interaction. However, studies in the emerging field of neuroeconomics cast doubt on this conclusion.

Recent findings show that people systematically reach different decisions and show different patterns of brain activation with computers in the exact same decision making tasks, for the exact same financial incentives, when compared to humans. Gallagher et al. [14] showed that when people played the rock-paper-scissors game with a human there was activation of the medial prefrontal cortex (MPFC), which had previously been shown to be involved in mentalizing (i.e., inferring of the other's beliefs, desires and intentions); however, no such activation occurred when people engaged with a computer that followed a predefined algorithm to make the choice. McCabe et al. [15] found a similar pattern when people played the trust game with humans in comparison to a probabilistic algorithm; Riedl et al. [16] further replicated this result with virtual humans, i.e., computers with three-dimensional virtual bodies and faces. In the prisoner's dilemma, Rilling et al. [17] and Krach et al. [18] showed that people tended to cooperate more with humans than computers and, once again, brain regions associated with mentalizing such as the MPFC, the rostral anterior cingulate cortex and the right temporo-parietal junction, were only activated with humans; in contrast, Kircher et al. [19] showed no difference in cooperation rates between humans and computers, despite reporting the usual increased brain activity with humans. In an influential paper, Sanfey et al. [20] showed that people were more willing to accept unfair offers in the ultimatum game from a computer than from a human. Moreover, their results revealed that the bilateral anterior insula–a region usually associated with the experience of negative emotions–showed higher activation when people received unfair offers from humans than from computers, thus suggesting that increased negative emotion explained the discrepancies in decision making behavior. Complementing this work, van't Wout et al. [21] showed that unfair offers in the ultimatum game led skin conductance–an autonomic index of affective state–to raise with humans but not with computers. In sum, these findings suggest that, in social decision making, people reach different decisions with computers when compared to humans.

From a theoretical perspective, there are two contrasting views regarding the social impact of computers on people's behavior. The "computers as social actors" theory [22], [23] argues that as long as machines displays social cues (e.g., nonverbal behavior) people will treat them in a fundamentally social manner. The argument is that people "mindlessly" treat computers that exhibit social traits like other people as a way to conserve cognitive effort and maximize response efficiency [24]. According to this theory, thus, computer agents that display emotions should impact humans in the same manner as people that show emotion. In contrast, Blascovich and colleagues [25], [26] argue that, everything else being equal, social influence will be greater the higher the perceived "agency" of the computer. Agency refers to people's theories of mind regarding the computer, i.e., the perceived sentience (e.g., attributions of consciousness, free will). According to this theory, thus, people will treat computers differently according to whether they believe it is being controlled by a computer algorithm (i.e., an *agent*) or a human (i.e., an *avatar*). The aforementioned findings in neureoeconomics would be in line with this latter theory.

The goal of this paper is, therefore, to study whether people's decisions when engaging with emotional virtual humans in social decision making tasks, for clear financial stakes, will change according to perceived agency, i.e., the belief about whether the virtual human is an agent or an avatar. To accomplish this we present two novel experiments where people engaged in a social dilemma (Experiment 1) and in negotiation (Experiment 2) with virtual humans that were described to be either agents or avatars. In line with Blascovich et al.'s social influence theory and the aforementioned findings in neuroeconomics, our general hypothesis was that:

Hypothesis 1. *The higher the perceived agency, the stronger the social effects of emotion expressions would be on people's decision making.*

II. Experiment 1

In this experiment participants engaged in a social dilemma with emotional virtual humans. Social dilemmas are situations where an individual gets a higher payoff by defecting rather than cooperating, regardless of what others in society do, yet all individuals end up receiving a lower payoff if all defect than if all cooperate [27]. In this experiment, participants engaged in the iterated prisoner's dilemma, a social dilemma commonly used to study emergence of cooperation. The prisoner's dilemma is a two-player game where the payoffs of each player depend on the simultaneous choice of both players. The payoff matrix for this task is shown in Table 1. The task represents a dilemma because the rational (i.e., utility-maximizing) choice for both players is to defect, which results in an outcome (mutual defection) that is worse than mutual cooperation. Participants played 20 rounds of this task. Moreover, following the approach by Kiesler, Waters and Sproull [28], the task was recast as an investment game.

TABLE I. Payoff Matrix for the Prisoner's Dilemma

		Virtual Human: Cooperation		*Virtual Human*: Defection	
Participant	Cooperation	VH:	6 pts	VH:	10 pts
		Partic:	6 pts	Partic:	0 pts
	Defection	VH:	0 pts	VH:	3 pts
		Partic:	10 pts	Partic:	3 pts

Researchers have argued that emotion expressions can signal others that one is willing to cooperate in a social exchange and people look for such cues before making a decision [29], [30], [31]. Similarly, emotion displays can also signal that one has competitive intentions [32], [33]. Accordingly, we [12], [34], [35] showed in a series of experiments that people would cooperate more with virtual humans that showed cooperative facial displays (e.g., joy after mutual cooperation) than virtual humans that showed competitive displays (e.g., joy after exploiting the participant). Following these findings, we defined the *expressively cooperative* counterpart (Table 2, top), which displays joy in mutual cooperation and regret when it exploits the participant, and the *expressively competitive* counterpart (Table 2, bottom), which displays joy when it exploits the participant and regret in mutual cooperation. The rationale for the cooperative counterpart is that joy after mutual cooperation signals an intention to cooperate, whereas regret after exploitation acknowledges the transgression; the rationale for the competitive counterpart is that joy after exploitation signals an intention to compete, whereas regret after mutual cooperation signals regret for missing the chance to exploit the participant.

TABLE II. The Virtual Human's Facial Displays

Expressively Cooperative		*Virtual Human*: Cooperation	*Virtual Human*: Defection
Participant	Cooperation	Joy	Regret
	Defection	Neutral	Neutral

Expressively Competitive		*Virtual Human*: Cooperation	*Virtual Human*: Defection
Participant	Cooperation	Regret	Joy
	Defection	Neutral	Neutral

A limitation of our earlier experiments was that perceived agency was not manipulated and, thus, it was not clear whether virtual humans were perceived as agents or avatars. For instance, virtual humans were always addressed by a name (e.g., "Ethan") and, participants were not explicitly instructed that they would be engaging with computer agents. In contrast, in the current experiment we carefully manipulated people's perceived agency. Agents were always referred to as "computer agents" and were described to the participants as "a computer program that was designed to make decisions just like other people". Avatars were described as "the players' visual representation in the game". Participants were asked to choose an avatar for themselves, of the same gender, and were informed that their avatar "would be visible to the other player" and that the they "would be able to control aspects of the avatar's behavior which would be visible to the other player, and vice-versa". Participants in both agent and avatar conditions chose an avatar for themselves. In avatar conditions, the counterpart's avatar was described to be controlled by another participant. In reality, participants always played with a computer program that followed the same strategy: tit-for-tat, starting with a defection. To make this deception believable, we further implemented a server that matched pairs of participants that were supposed to engage with other participants; participants would then proceed in lockstep throughout the task but the responses they would see always followed the tit-for-tat strategy. Participants were also made to believe they were engaging with a participant of the same gender when, in fact, this might have not been the case. Lastly, participants were told that the identities of other participants would be concealed and the software always referred to the human counterpart as "anonymous". Following the discussion in the Introduction, our hypothesis was that:

> ***Hypothesis 2.*** *People would cooperate more with the cooperative than the competitive virtual human but, this difference would only be significant with avatars.*

A. *Design*

The experiment followed a 2 × 2 between-participants factorial design: *Emotion Displays* (Cooperative vs. Competitive) × *Agency* (Agent vs. Avatar). We used the same emotion facial displays that were validated and used in our earlier experiments [12], [34], [35]: joy was expressed through a smile and contraction of the corrugator supercilii (eyes); regret was expressed through lowering of the zygomaticus, light blushing, head bowing and gaze aversion. Male and female avatars used in the experiment are shown in Figure 1.

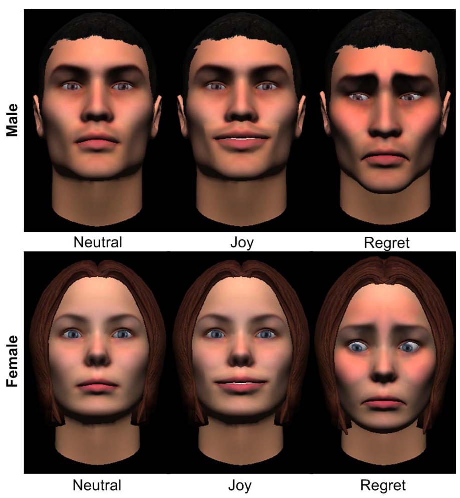

Fig. 1. The emotion facial displays used in the experiment.

Our main dependent variable was cooperation rate, i.e., the number of times participants cooperated over all rounds. To validate that participants were correctly perceiving some virtual humans as agents and others as avatars we asked them, after the task was completed, to rate the virtual human according to the following pairs of adjectives on a 7-point scale (e.g., for Fake-Natural, 1 corresponded to Fake and 7 to Natural): Robot like-Human like; Fake-Natural; Unconscious-Conscious; Artificial-Lifelike; Stagnant-Lively; Mechanical-Organic; Inert-Interactive; Apathetic-Responsive; and, Computer-Human. The selection of these adjectives was based on existent scales pertaining to anthropomorphism [36], the "uncanny valley" effect (e.g. [37]) and the experience of social presence in virtual environments (e.g., [38]).

One-hundred and twenty six participants were recruited at USC's Marshall School of Business. This resulted in approximately 30 participants per condition. Regarding gender, 69.7% were males. Age distribution was as follows: 21 years and Under, 70.6%; 22 to 34 years, 29.4%. Most participants were undergraduate students (95.8%) majoring in Business-related courses and with citizenship from the United States (81.5%). The incentive to participate followed standard practice in experimental economics [39]: first, participants were given school credit for their participation; second, with respect to their goal in the task, participants were instructed to earn as many points as possible, as the total amount of points would increase their chances of winning a lottery for $100. Upon completion of the experiment participants were verbally debriefed about the deception pertaining to the avatar conditions.

B. *Results*

Participants that did not experience both joy and regret with the counterpart[1] – i.e., our experimental manipulation – were excluded from analysis (though keeping them would lead to the same pattern of results). After exclusion, 84 participants remained for analysis.

Regarding the agency manipulation check, the nine adjective classification questions were highly correlated (Cronbach α = .972) and, thus, were averaged into a single measure we called anthropomorphism. We then ran an Emotion Displays × Agency ANOVA which revealed no main effect of Emotion Displays, $F(1, 80) = 1.13$, $p = .291$, but, as expected, confirmed a main effect of Agency, $F(1, 80) = 4.48$, $p = .037$, partial $\eta^2 = .053$: people perceived avatars ($M = 4.87$, $SD = 1.54$) to be more anthropomorphic than agents ($M = 4.12$, $SD = 1.64$). The Emotion Displays × Agency interaction was not significant, $F(1, 80) = .541$, $p = .464$.

Regarding cooperation rate, the means and standard errors are shown in Figure 2. To test our Hypothesis 2, we split the data across Agency and ran independent t tests to compare cooperation rates between cooperative and competitive virtual humans. This analysis revealed that, for agents, people cooperated more with cooperative ($M = .64$, $SD = .26$) than competitive agents ($M = .54$, $SD = .30$) but this result did not reach significance, $t(37) = 1.12$, $p = .269$, $r = .181$. For avatars, people cooperated more with cooperative ($M = .73$, $SD = .26$) than competitive avatars ($M = .55$, $SD = .28$) and this result was significant, $t(43) = 2.31$, $p = .026$, $r = .332$.

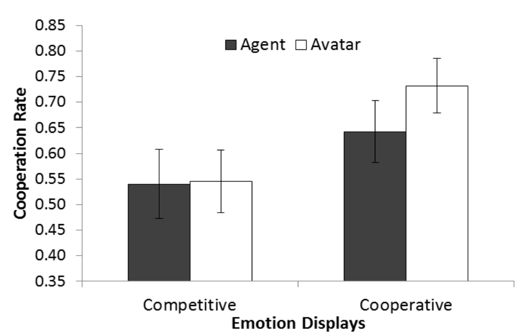

Fig. 2. Means (and standard errors) for cooperation rate.

[1] Notice this paradigm did not guarantee participants would experience all outcomes in the prisoner's dilemma task.

C. Discussion

The results confirmed that people reach different decisions when engaged with computer algorithms that honestly portray themselves as computers compared to if they portray themselves as human. Merely this belief – even though the financial incentives and the virtual human's appearance, decisions, and expressions were identical – had a powerful effect. In support of Hypothesis 2, emotion displays shaped participants' willingness to cooperate, but these effects were only significant, and the effect sizes much larger (r = .332 vs. r = .181) when playing against a presumed human opponent. Strikingly, as cooperative emotions promoted greater cooperation rates and thus greater individual rewards, participants were able to earn more money when they were deceived about the true nature of the virtual human.

III. Experiment 2

In this experiment participants engaged in negotiation, a domain inherently different from social dilemmas [40], with emotional virtual humans. According to Pruitt and Carnevale [41], negotiation is "a discussion among two or more parties aimed at reaching agreement when there is a perceived divergence of interest". In this experiment, similarly to Van Kleef et al.'s approach [5], people engaged in a multi-issue negotiation assuming the role of a seller of a phone company whose goal was to negotiate three issues: the price, the warranty period and the duration of the service contract of the phones. Each issue had 9 levels, being the highest level the most valuable for the participant, and the lowest level the least valuable. Level 1 on price ($110) yielded 0 points and level 9 ($150) yielded 400 points (i.e., each level corresponded to a 50 point increment). Level 1 on warranty (9 months) yielded 0 points and level 9 (1 month) yielded 120 points (i.e., each level corresponded to a 15 point increment). For duration of service contract, level 1 (9 months) yielded 0 points, and level 9 (1 month) yielded 240 points (i.e., each level corresponded to a 30 point increment). It was pointed out to the participant that the best deal was, thus, 9-9-9 for a total outcome of 760 points (400 + 120 + 240). The participant was also told that the counterpart had a different payoff table which was not known. The negotiation would proceed according to the alternating offers protocol, being the counterpart the first to offer, and until someone accepted the other's offer or "time expired"; in reality, if no agreement had been reached, the task would always terminate in round 6.

Recently, researchers began looking at the impact of emotion displays on negotiation outcome (e.g., [1]) and a finding that is relevant to this work is that people concede more when facing an angry than a neutral counterpart [5], [42]. The argument is that people infer the angry counterpart to have high aspirations and, so as to avoid costly impasse, are forced to lower their demand. Recently, we [13] also showed that people conceded more to angry than happy virtual humans. However, once again, there was some ambiguity regarding the virtual human's agency (e.g., they were referred to by a name, such as "Ethan"), and the experiment did not manipulate perceived agency; thus, it was not clear whether participants perceived the virtual humans to be avatars or agents. In contrast, in this experiment we explicitly manipulated perception of agency. Virtual humans were described in a similar manner as in Experiment 1 (e.g., the agent was always referred to as "computer agent" and the avatar was always described as "anonymous"). In reality, participants always engaged with a computer program that followed a scripted strategy. When participants engaged with avatars, we used a server to implement the deception. In this case, the server would synchronize the participants at the beginning of the task and, from that point onward, the scripted strategy would be played. Participants were also made to believe they were engaging with another participant of the same gender, even though that might have not been the case. Following the discussion in the Introduction, our hypothesis was that:

> ***Hypothesis 3.*** *People would concede more to angry than neutral virtual humans but, this difference would only be significant with avatars.*

A. Design

The experiment followed a 2 × 2 factorial between-participants design: *Emotion Displays* (Neutral vs. Anger) × *Agency* (Agent vs. Avatar). We used the same emotion facial displays that were validated and used in our earlier experiment [13]. One male and one female avatars are shown in Figure 3. Emotion displays would be shown after the participant had made an offer in rounds 1, 3 and 5. Regarding strategy, independently of the agency condition, participants always saw the same fixed sequence of offers: 2-3-2, 2-3-3, 2-4-3, 3-4-3, 3-4-4, and 4-4-4. This pattern had been argued before to strike a good balance between cooperation and competition [5].

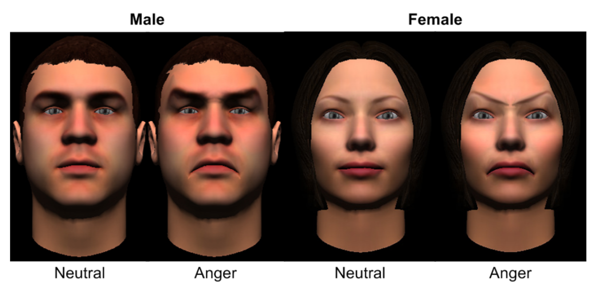

Fig. 3. The emotion facial displays used in the experiment.

Regarding measures, our main dependent variable was *demand difference*, i.e., the difference in demand level between round 1 (initial offer) and round 6 (final offer). To calculate demand level, the number of points demanded in each round was summed across all issues of price, warranty and service. Demand difference was then calculated by subtracting demand level in round 1 (first offer) and demand level in round 6 (last offer). To validate that participants were correctly perceiving some virtual humans as agents and others as avatars we asked them to rate the virtual human on the same adjective pairs as in Experiment 1.

Seventy-eight participants were recruited at the paid pool at USC's Marshall School of Business. This resulted in approximately 20 participants per condition. Regarding gender, 45.8% were males. Age distribution was as follows: 21 years and Under, 52.8%; 22 to 34 years, 47.2%. Most participants were undergraduate (63.9%) and graduate (34.7%) students majoring in diverse fields and mostly with citizenship from the United States (59.7%) and India (27.8%). The incentive to participate followed standard practice in experimental economics [39]: first, participants were paid \$20 for their participation; second, with respect to their goal in the task, participants were instructed to earn as many points as possible, as the total amount of points would increase their chances of winning a lottery for \$100. Upon completion of the experiment participants were verbally debriefed about the deception pertaining to the avatar conditions.

B. Results

Participants that accepted the virtual human's first offer or whose first offer was accepted by the virtual human did not see any emotion expression – i.e., our experimental manipulation – and, thus, were excluded from analysis (though keeping them would lead to the same pattern of results). After exclusion, 72 participants remained for analysis.

Regarding the agency manipulation check, the nine adjective classification questions were highly correlated (Cronbach α = .952) and, thus, were averaged into a single measure we called anthropomorphism. We then ran an Emotion Displays × Agency ANOVA which revealed no main effect of Emotion Displays, $F(1, 68) = 2.96$, $p = .090$, but, as expected, confirmed a main effect of Agency, $F(1, 68) = 9.87$, $p = .002$, partial $\eta^2 = .127$: people perceived avatars ($M = 3.84$, $SD = 1.29$) to be more anthropomorphic than agents ($M = 2.98$, $SD = 1.38$). The Emotion Displays × Agency interaction was not significant, $F(1, 68) = 2.45$, $p = .123$.

Regarding demand difference, the means and standard errors are shown in Figure 4. To test Hypothesis 3, we split the data across Agency and ran independent *t* tests to compare demand difference between angry and neutral virtual humans. This analysis revealed that, for agents, demand difference was higher with angry agents ($M = 166.75$, $SD = 160.19$) than neutral agents ($M = 157.25$, $SD = 127.38$) but this result was not significant, $t(38) = -.208$, $p = .837$, $r = .034$. For avatars, demand difference was higher with angry avatars ($M = 286.67$, $SD = 218.55$) than neutral avatars ($M = 101.50$, $SD = 111.57$) and this results was significant, $t(30) = -3.182$, $p = .003$, $r = .502$.

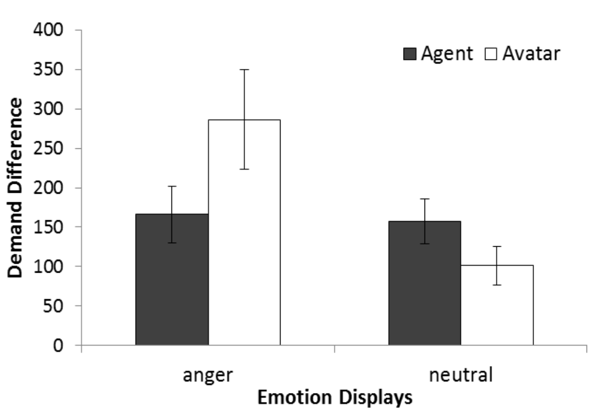

Fig. 4. Means (and standard errors) for demand difference.

C. Discussion

The results showed, once again, that people behaved differently when they perceived virtual humans to be agents when compared to avatars. Virtual humans were unsuccessful in influencing participant behavior when they honestly portrayed themselves as computers ($r = .034$); however they had a large effect on behavior ($r = .502$) when they portrayed themselves as a human, in which case people conceded significantly more to an angry than a neutral avatar. Overall, thus, the results supported Hypothesis 3.

IV. General Discussion

Our results show that the belief about whether a computer or a human is controlling the emotion expressions of a virtual human can significantly impact people's decision making behavior. In Experiment 1, we had people engage in a social dilemma with virtual humans that showed either cooperative (e.g., joy after mutual cooperation) or competitive displays (e.g., joy after exploiting the participant). The results showed that people always cooperated more with the cooperative than the competitive virtual humans but, this difference was only significant when they believed they were engaging with avatars. In Experiment 2, we had people engage in negotiation with virtual humans that showed either an angry or neutral expression. The results showed that people conceded more to the angry than the neutral virtual human but, once again, this difference was only significant with avatars. Overall, and in support of Hypothesis 1, the results suggest that the effects of emotions expressions on people's decisions are stronger the higher the perceived agency of the virtual human.

These results contrast with the "computers as social actors" theory [22], [23] which argues people treat computers that display social cues in the same manner as people and, rather, are compatible with Blascovisch et al.'s view [25], [26] that social influence will be greater the higher people's attributions of a mind to the computer, i.e., the perceived sentience (e.g., consciousness or free will). The results are also compatible with recent findings in neuroeconomics which suggest that brain regions usually associated with mentalizing tend to show higher activation patterns in decision making tasks when people believe they are engaging with humans rather than computers. Effectively, we had argued elsewhere [35], [34] that a key for the social effects of emotion expressions is the information people retrieve from such displays about the other's beliefs, desires and intentions. In this sense, a higher activation of the mentalizing brain regions with humans might have meant people tried harder to infer the human's mental states from their emotion displays, which then led to increased effects when compared to computers.

The agency effect reported in this paper seems to reflect people's current suspicion about the ability of a machine to "have a mind", i.e., a mind that is worthy of mentalizing as is the mind of a human. Future work should, then, systematically explore which cognitive abilities need to be simulated in our agents so that we cross the threshold above which people cease to distinguish between humans and computers, at least in the context of social decision making tasks with clear financial stakes.

ACKNOWLEDGMENT

This research was supported in part by the Air Force Office of Scientific Research (AFOSR) and grant 1211064 from the NSF. The content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

[1] G. Van Kleef, C. De Dreu, and A. Manstead, "An interpersonal approach to emotion in social decision making: The emotions as social information model," Ad Exp Soc Psychol vol. 42, pp. 45-96, 2010.

[2] N. Frijda, and B. Mesquita. "The social roles and functions of emotions," in Emotion and culture: Empirical studies of mutual influence, S. Kitayama and H. Markus, Eds. Washington, DC: American Psychological Association, 1994 , pp. 51–87.

[3] D. Keltner, and J. Haidt, "Social functions of emotions at four levels of analysis," Cognition and Emotion, vol. 13, pp. 505–521, 1999.

[4] D. Keltner, and A. Kring, "Emotion, social function, and psychopathology," Rev Gen Psychol, vol. 2, pp. 320–342, 1998.

[5] G. Van Kleef, C. De Dreu, and A. Manstead, "The interpersonal effects of anger and happiness in negotiations," J Pers Soc Psychol, vol. 86, pp. 57-76, 2004.

[6] G. Van Kleef, C. De Dreu, and A. Manstead, "Supplication and appeasement in negotiation: The interpersonal effects of disappointment, worry, guilt, and regret," J Pers Soc Psychol, vol. 91, pp. 124-142, 2006.

[7] E. Krumhuber, A. Manstead, and A. Kappas, "Facial dynamics as indicators of trustworthiness and cooperative behavior," Emotion, vol. 7, pp. 730-735, 2007.

[8] E. van Dijk, G. Van Kleef, W. Steinel, and I. Van Beest, "A social functional approach to emotions in bargaining: When communicating anger pays and when it backfires," J Pers Soc Psychol, vol. 94, pp. 600-614, 2008.

[9] M. Wubben, D. De Cremer, and E. van Dijk, "When and how communicated guilt affects contributions in public good dilemmas," J Exp Soc Psychol, vol. 45, pp. 15-23, 2009.

[10] R. Friedman, C. Anderson, J. Brett, M. Olekalns, N. Goates, and C. Lisco, "The positive and negative effects of anger on dispute resolution: Evidence from electronically mediated disputes," J Appl Psychol, vol. 89, pp. 369-376, 2004.

[11] B. Parkinson, and G. Simons, "Affecting others: Social appraisal and emotion contagion in everyday decision making," Pers Soc Psychol B, vol. 35, pp. 1071-1084, 2009.

[12] C. de Melo, P. Carnevale, and J. Gratch, "The impact of emotion displays in embodied agents on emergence of cooperation with people," Presence: Teleoperators and Virtual Environments, vol. 20, pp. 449-465, 2012.

[13] C. de Melo, P. Carnevale, and J. Gratch, "The effect of expression of anger and happiness in computer agents on negotiations with humans," in Proceedings of Autonomous Agents and Multiagent Systems (AAMAS), 2011.

[14] H. Gallagher, J. Anthony, A. Roepstorff, and C. Frith, "Imaging the intentional stance in a competitive game," NeuroImage, vol. 16, pp. 814-821, 2002.

[15] K. McCabe, D. Houser, L. Ryan, V. Smith, and T. Trouard, "A functional imaging study of cooperation in two-person reciprocal exchange," PNAS, vol. 98, pp.11832-11835, 2001.

[16] R. Riedl, P. Moht, P. Kenning, F. Davis, and H. Heekeren, "Trusting humans and avatars: Behavioral and neural evidence," in Proceedings of the 32nd Int. Conference on Information Systems, 2011.

[17] J. Rilling, D. Gutman, T. Zeh, G. Pagnoni, G. Berns, and C. Kilts, "A neural basis for social cooperation," Neuron, vol. 35, pp.395-405, 2002.

[18] S. Krach et al., "Can machines think? Interaction and perspective taking with robots investigated via fMRI," PLoS ONE, vol.3, pp.1-11, 2008.

[19] T. Kircher, et al., "Online mentalising investigated with functional MRI," Neurosci Lett, vol. 454, pp. 176-181, 2009.

[20] A. Sanfey, J. Rilling, J. Aronson, L. Nystrom, and J. Cohen, "The neural basis of economic decision-making in the ultimatum game," Science, vol.300, pp.1755-1758, 2003.

[21] M. van't Wout, R. Kahn, A. Sanfey, and A. Aleman, "Affective state and decision-making in the ultimatum game," Exp Brain Res, vol.169, pp.564-568, 2006.

[22] C. Nass, J. Steuer, and E. Tauber, "Computers are social actors," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1994.

[23] B. Reeves, and C. Nass. The media equation: How people treat computers, television, and new media like real people and places. New York, NY: Cambridge University Press.

[24] C. Nass, and Y. Moon, "Machines and mindlessness: Social responses to computers," Journal of Social Issues, vol.56, pp.81-103, 2000.

[25] J. Blascovich et al., "Immersive virtual environment technology as a methodological tool for social psychology," Psychological Inquiry, vol.13, pp.103-124, 2002.

[26] J. Blascovich, and C. McCall, "Social influence in virtual environments," in The Oxford Handbook of Media Psychology, K. Dill, Ed., New York, NY: Oxford University Press, 2013, pp.305-315.

[27] R. Dawes, "Social dilemmas," Annu Rev Psychol, vol. 31, pp.169-193, 1980.

[28] S. Kiesler, K. Waters, and L. Sproull, "A prisoner's dilemma experiment on cooperation with human-like computers," J Pers Soc Psychol, vol.70, pp.47-65, 1996.

[29] R. Frank, Passions within reason. New York, NY: Norton, 1988.

[30] R. Nesse, "Evolutionary explanations of emotions," Human Nature, vol.1, pp.261-289, 1990.

[31] R. Trivers, "The evolution of reciprocal altruism," Q Rev Biol, vol.46, pp.35-57, 1971.

[32] D. Matsumoto, and B. Willingham, "The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic Games," J Pers Soc Psychol, vol.91, pp.568–581, 2006.

[33] D. Matsumoto, N. Haan, Y. Gary, P. Theodorou, and C. Cooke-Carney, "Preschoolers' moral actions and emotions in prisoner's dilemma," Dev Psychol, vol.22, pp.663-670, 1986.

[34] C. de Melo, P. Carnevale, and J. Gratch, "Reverse appraisal: The importance of appraisals for the effect of emotion displays on people's decision-making in a social dilemma." in Proceedings of 34th Annual Meeting of the Cognitive Science Society, 2012.

[35] C. de Melo, The interpersonal effect of emotion in decision-making and social dilemmas, Ph.D. diss., Department of Computer Science, University of Southern California, Los Angeles, CA, 2012.

[36] N. Epley, A. Waytz, and J. Cacioppo, "On seeing human: A three-factor theory of anthropomorphism," Psych Rev, vol.114, pp.864-886, 2007.

[37] C. Bartneck, D. Kulic, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," Int J Soc Rob, vol.1, pp.71-81, 2009.

[38] C. Harms, and F. Biocca, "Internal consistency and reliability of the networked minds measure of social presence," in Proceedings of 7th Annual Workshop: Presence, 2004.

[39] R. Hertwig, and A. Ortmann, "Experimental practices in economics: A methodological challenge for psychologists?," Behav Brain Sci, vol.24, pp.383-451, 2001.

[40] D. Pruitt, and M. Kimmel, "Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future," Annu Rev Psychol, vol.28, pp.363-392, 1977.

[41] D. Pruitt, and P. Carnevale, Negotiation in social conflict. Pacific Grove, CA: Brooks/Cole, 1993.

[42] M. Sinaceur, and L. Tiedens, "Get mad and get more than even: When and why anger expression is effective in negotiations," J Exp Soc Psychol, vol.42, pp.314-322, 2006.