

People Don't Feel Guilty About Exploiting Machines

CELSO DE MELO, University of Southern California

STACY MARSELLA, Northeastern University

JONATHAN GRATCH, University of Southern California

Guilt and envy play an important role in social interaction. Guilt occurs when individuals cause harm to others or break social norms. Envy occurs when individuals compare themselves unfavorably to others and desire to benefit from the others' advantage. In both cases, these emotions motivate people to act and change the status quo: following guilt, people try to make amends for the perceived transgression and, following envy, people try to harm envied others. In this paper, we present two experiments that study participants' experience of guilt and envy when engaging in social decision making with machines and humans. The results showed that, though experiencing the same level of envy, people felt considerably less guilt with machines than with humans. These effects occurred both with subjective and behavioral measures of guilt and envy, and in three different economic games: public goods, ultimatum, and dictator game. This poses an important challenge for human-computer interaction because, as shown here, it leads people to systematically exploit machines, when compared to humans. We discuss theoretical and practical implications for the design of human-machine interaction systems that hope to achieve the kind of efficiency – cooperation, fairness, reciprocity, etc. – we see in human-human interaction.

• Human-centered computing → Human computer interaction (HCI) → HCI theory, concepts and models • Human-centered computing → Human computer interaction (HCI) → Empirical studies in HCI.

Additional Key Words and Phrases: Guilt, envy, decision making, humans vs. machines

1. INTRODUCTION

The last few decades have seen growing interest in the development of machines (software agents or robots) that can engage in decision making with humans [de Melo et al. 2011, 2012], [Lin and Kraus 2010], [Jennings et al. 2001]. Moreover, companies are investing billions of dollars creating computational systems that make autonomous decisions in various domains, such as finance, healthcare, and public policy [Davenport and Harris 2005]. As machines become pervasive in all aspects of personal, social, and professional life, it is important we develop theories that explain human decision making with machines. The challenge is that, despite evidence that people can treat machines like humans in social settings [Nass and Moon 2000], [Nass et al. 1996, 1997, 1999, 2000], [Reeves and Nass 1996], recent research shows that there are still important differences in the way people behave with machines, when compared to humans. This evidence demonstrates that people systematically favor humans to machines when engaging in social decision making [Gallagher et al. 2002], [Kircher et al. 2009], [Krach et al. 2008], [McCabe et al. 2001], [Rilling et al. 2002], [Sanfey et al. 2003]. Here we show that one of the reasons this bias occurs is that people experience emotion differently with machines than with humans.

This work is supported by the National Science Foundation, under grant SES-0836004, and the Air Force Office of Scientific Research, under grant FA9550-09-1-0507. The content does not necessarily reflect the position or the policy of any Government, and no official endorsement should be inferred.

Author's addresses: C. de Melo, Institute for Creative Technologies, University of Southern California; Stacy Marsella, Northeastern University; J. Gratch, Institute for Creative Technologies, University of Southern California.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

Copyright © ACM 2015 1073-0516/2015/MonthOfPublication - ArticleNumber \$15.00

DOI:<http://dx.doi.org/10.1145/0000000.0000000>

Specifically, we demonstrate that, despite experiencing as much envy with machines as with humans, people experience less guilt when exploiting machines. To the best of our knowledge, this is the first time clear evidence is shown that increased guilt is critical in explaining people's behavioral differences with machines, when compared to humans. This result has important theoretical implications for our understanding of the nature of people's interaction with machines, as well as practical implications for the design of human-machine interaction systems that can achieve the same level of cooperation, fairness, reciprocity, etc., we see in human-human interaction.

1.1 Treating Machines in a Social Manner

Nass and colleagues [Nass and Moon 2000], [Nass et al. 1996, 1997, 1999, 2000], [Reeves and Nass 1996] were among the first to advance a general theory for human-machine interaction. The theory's main tenet is that to the extent that machines display social cues (e.g., interactivity, verbal and nonverbal behavior), people will treat them in a fundamentally social manner and apply the same rules they use when interacting with other people. Some of the studies supporting this theory show that people build rapport with machines [Gratch et al. 2007], react to emotion expressed by machines [de Melo et al. 2014], demonstrate politeness to machines [Nass et al. 1999], form positive impressions of machines that are perceived as team members [Nass et al. 1996], and apply stereotypes to machines [Nass et al. 1997, 2000].

These initial findings were so promising that Reeves and Nass [1996] actually proposed that it was possible to replicate any finding in human-human interaction with computers:

“Findings and experimental methods from the social sciences can be applied directly to human-media interaction. It is possible to take a psychology research paper about how people respond to other people, replace the word ‘human’ with the word ‘computer’ and get the same results.” [Reeves and Nass 1996; pg. 28]

Thus, a strict interpretation of the theory suggests that people, first, should demonstrate social considerations when engaging with machines in economic settings; and, second, they should make offers to machines at the same level as with humans. Nevertheless, despite the initial evidence suggesting that machines are treated like humans, recent research shows that people still make important distinctions between machines and humans.

1.2 People Act Differently with Machines

Several studies have now demonstrated that people behave differently with machines, in comparison to humans [Gallagher et al. 2002], [Kircher et al. 2009], [Krach et al. 2008], [McCabe et al. 2001], [Rilling et al. 2002], [Sanfey et al. 2003], [van't Wout et al. 2006]. This research shows that people can reach different decisions and show different patterns of brain activation with machines in the exact same decision making tasks, for the exact same financial incentives, when compared to humans. Gallagher et al. [2002] showed that when people played the rock-paper-scissors game with a human there was activation of the medial prefrontal cortex, a region of the brain that had previously been implicated in mentalizing (i.e., inferring of other's beliefs, desires and intentions); however, no such activation occurred when people engaged with a machine that followed a known predefined algorithm. McCabe et al. [2001] found a similar pattern when people played the trust game with humans vs. computers, and Kircher et al. [2009], Krach et al. [2008], and Rilling et al. [2002]

replicated this finding in the prisoner's dilemma. The evidence, thus, suggests that people spent less effort inferring mental states with machines than with humans. These findings are compatible with research that shows that people perceive less mind in machines than in humans [Gray et al. 2007], [Waytz et al. 2010]. Denying mind to others or perceiving inferior mental ability in others, in turn, is known to lead to discrimination [Haslam 2006] and, thus, may explain why people favored humans to machines. In the context of human-machine interaction, Blascovich et al. [2002] also propose that machines are less likely to influence humans, the lower the perceived mental ability in machines.

Building on some of these earlier findings, Sanfey et al. [2003] further showed that, when receiving unfair offers in the ultimatum game, people showed stronger activation of the bilateral anterior insula – a region associated with the experience of negative emotions – when engaging with humans, when compared to machines. Van't Wout et al. [2006] replicated this finding in the ultimatum game and presented evidence that people show less arousal, as measured by skin conductance, with machines than with humans. In non-economic tasks, human-computer interaction researchers also replicated the finding that people experience higher arousal with humans than with machines [Lim and Reeves 2010], [Mandryk et al. 2006], [Okita et al. 2007], [Ravaja 2009]. This evidence, therefore, suggests that people experience emotion differently when engaging with machines, when compared to humans. In this paper, we extend the existent literature by studying the specific role of guilt and envy in explaining this difference in behavior when people engage with human vs. machine counterparts.

1.3 Guilt and Envy in Social Interaction

Research shows that the experience of emotion has a profound effect on people's behavior in social situations [Damasio 1994], [Lerner et al. 2015], [Loewenstein and Lerner 2003], [Pillutla and Murnighan, 1995]. Guilt and envy, in particular, have important interpersonal consequences. The experience of guilt is associated with a perceived transgression, in particular, when the individual feels that he or she has caused harm to others or has violated some expectation or norm [Leith and Baumeister 1998], [Lewis 1971], [Tangney 1990, 1995, 1999]. Guilt, thus, tends to produce outwardly focused behavior aimed at repairing the transgression and changing the current state-of-affairs [Carlsmith and Gross 1969], [Freedman et al. 1967], [Ketelaar and Au 2003], [Van Kleef et al. 2006]. In contrast, the experience of envy occurs when individuals compare themselves unfavorably to others and, consequently, desire to benefit from what envied others have [Duffy et al. 2012], [Smith and Kim 2007]. Envy can lead, on the one hand, to positive consequences when it motivates individuals to work harder and attempt self-improvement; on the other hand, it can lead to negative consequences when it leads individuals to take action to undermine the envied person's advantage.

Guilt and envy also have a strong influence on decision making behavior. In such settings, Fehr and Schmidt [1999] propose an operationalization of guilt and envy that is useful for our purposes. Accordingly, guilt is said to occur when one perceives one was given an unfair advantage over others; in contrast, envy is experienced when one feels others have been given an unfair outcome advantage. To capture these intuitions, they proposed an economic model that is now widely known: the inequity aversion model [Fehr and Schmidt 1999]. The idea behind this model is that people are inequity averse, i.e., people dislike outcomes where others are better off than

them (envy) or outcomes where they are better off than others (guilt). Thus, they propose the following utility function for dyadic interactions:

$$U_k(x_k, x_l) = x_k - \alpha_k \max\{x_l - x_k, 0\} - \beta_k \max\{x_k - x_l, 0\} \quad (1)$$

with $0 \leq \beta_i \leq \alpha_i$ and $\beta_i \leq 1$ and where x_i , $i \in \{k, l\}$, are the payoffs for the players. Notice that the parameter α captures aversion to disadvantageous inequality (envy), whereas the parameter β captures aversion to advantageous inequality (guilt). This simple utility function has successfully modeled people's behavior in many economic settings, including the ultimatum and dictator games [Fehr and Schmidt 2006]. This model has also been used by others to study envy and guilt in human-human interaction (e.g., [Krajbich et al. 2009]).

Research shows that emotion is also very important in human-machine interaction [Marsella et al. 2010], [Picard 1997]. In this paper, we are specifically interested in people's experience of guilt and envy with machines. To study this, we use self-reported measures and Fehr and Schmidt's operationalization of these emotions. As discussed above, we expected people would experience emotion differently with machines, when compared to humans; in particular, our initial expectation was that people would experience more guilt and envy with humans than with machines.

1.4 Overview of Approach

To study people's experience of guilt and envy in human-machine interaction, we conducted two experiments where participants engaged in standard decision games with machine and human counterparts. These decision games capture prototypical social scenarios in which participants make decisions that have consequences for themselves and others [Pruitt and Kimmel 1977]. The decisions, moreover, had financial consequences, i.e., participants could earn real money according to their behavior. Money, in this case, was just a proxy for resources players care about in real life. Studying behavior in these standard settings is ideal for understanding the mechanisms underlying people's decision making, such as other-regarding preferences or cognitive biases. The critical comparison, thus, was between people's behavior and emotion experience with machines vs. humans.

In Experiment 1, participants engaged in the public goods game [Kollock 1998]. In this game there are two or more players who are given an initial endowment of money and have to simultaneously decide how much to contribute to a shared pool representing some public good. A public good is a shared resource (e.g., public television or city parks) from which all may benefit, regardless of whether they have helped provide for the good. The game, therefore, is an example of a social dilemma [Dawes 1980]. In social dilemmas, an individual always 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. Thus, the standard prediction in this game is for people to allocate nothing to the shared pool. Nevertheless, people often allocate a non-zero amount, especially when the number of players is low [Kollock 1998], [Isaac and Walter 1988], [Fehr and Gächter 2000]. Thus, following the previous discussion, we expected people to allocate a non-zero amount of money to the shared pool, but we expected this amount to be higher with humans than with machines.

In Experiment 2, we expanded on the previous experiment by considering two other decision tasks: the ultimatum and the dictator games. In the ultimatum game

[Güth et al. 1982], there are two players: a proposer and a responder. The proposer is given an initial endowment of money and has to decide how much to offer to the responder. Then, the responder has to make a decision: if the offer is accepted, both players get the proposed allocation; if the offer is rejected, however, no one gets anything. The standard rational prediction is that the proposer should offer the minimum non-zero amount, as the responder will always prefer to have something to nothing. In practice, people usually offer 40 to 50 percent of the initial endowment and low offers (about 20 percent of the endowment) are usually rejected [Henrich et al. 2001]. A concern for fairness is usually argued to explain this behavior. The dictator game [Forsythe et al. 1994] is similar to the ultimatum game, except that the responder doesn't get to make a decision and must accept whatever is offered by the proposer. This game, therefore, removes strategic considerations from the equation, i.e., the proposer doesn't have to fear that the offer will be rejected. Accordingly, the rational prediction is for the proposer to keep all the money and offer nothing. Nevertheless, even in this case, people offer an average of 10 to 25 percent of the initial endowment and the modal offer is 50 percent. Decisions in this game, thus, have been argued to reflect altruism. As in the previous experiment, we expected people to offer a non-zero amount of money to their counterparts, though we expected this amount to be higher with humans than with machines.

We used subjective and behavioral variables to measure the experience of guilt and envy. In Experiment 1, upon completion of the task, we asked participants to self-report their experience of guilt and envy. Subjective measures have the advantage of being easy to administer and are a good starting point for our research objectives. However, researchers [Bertrand and Mullainathan 2001], [Jahedi and Méndez 2014] have criticized subjective measures for being susceptible to (1) memory effects, since they are not in real-time, (2) presentation and image-preservation effects, since participants may provide socially desirable answers rather than reveal their true attitudes, and (3) framing and ordering effects, since the wording and order of the questions can influence people's answers. Moreover, subjective measures may not be able to capture phenomena that occur automatically or subconsciously. Thus, in Experiment 2, we resorted to Fehr and Schmidt's [1999] inequity aversion model to measure guilt and envy behaviorally. We used people's decisions in the ultimatum and dictator games to estimate their α (envy) and β (guilt) parameters. The question was whether people would have the same α and β parameters with machines as they did with humans. Our expectation, as discussed above, was that people would experience more guilt and envy, as measured both by subjective and objective measures, with humans than with machines.

2. EXPERIMENT 1

2.1 Method

2.1.1. Design. Participants engaged in the 2-player public goods game with human and computer counterparts. Each player was given an initial endowment of 20 lottery tickets. Participants engaged in 8 rounds of the game with the same counterpart. The tickets had real financial consequences as they would enter a lottery for \$30. Participants were instructed that the higher the amount of tickets they earned, the higher were their chances of winning the lottery. Participants were also instructed that computers would enter the lottery and, if one of them won, then no one else could win the lottery. Participants could allocate between 0 to 20 tickets to the shared pool. Tickets in this pool would then be multiplied by 1.5 and split equally by both players. So, for instance, if both players allocated all their tickets to the pool,

then each one would get, at the end, $40 \times 1.5 / 2 = 30$ tickets. We quizzed participants on these instructions before starting the task.

Participants engaged in a between-participants design with either human or computer counterparts. Humans were described as “other participants” and computers were “software programs that made decisions just like real people”. Independently of the condition, however, participants always engaged with the same computer script¹. To make this manipulation believable, we had people connect to a fictitious server before starting the task, for the purposes of “being matched with other players”. Connecting to this server took approximately 30-45 seconds. The counterpart script always followed a tit-for-tat strategy. Tit-for-tat has been shown to promote cooperation in other economic games [Axelrod and Hamilton 1981]. In our case, the counterpart always began by offering 18 tickets (90% of the initial endowment); in subsequent rounds, the counterpart offered whatever the participant offered in the previous round, plus some random variation (plus or minus 1 or 2 tickets). Upon completion of the task, we debriefed participants about this deception and the rationale for using it.

To minimize social pressure [Hoffman et al. 1994], participants’ identities were kept fully anonymous both to other participants and to the experimenters. To preserve anonymity between participants, human counterparts were referred to as “anonymous” and we never collected any information that could identify the participants. Computers were referred to as “computer agents”. To preserve anonymity with respect to the experimenter, we relied on Amazon Mechanical Turk – the online pool we used to recruit participants. When interacting with participants from this pool, researchers are never able to identify the participants, unless they explicitly ask for information that may serve to identify them (e.g., name or photo), which we did not. A screenshot of the software is shown in Figure 1.

2.1.2. Dependent Variables. The main behavioral variable was the allocation to the shared pool, averaged across all rounds. After completing the task, to measure the experience of guilt and envy we asked the following classification questions on a 7-point scale (1, *not at all*, to 7, *very much*):

- (1) Were you concerned with not taking advantage of the other party?
- (2) Did you feel guilty when the outcome favored you?
- (3) Were you comfortable with outcomes that favored you over the other party?
- (4) Were you concerned with not being taken advantage of?
- (5) Did you feel envious when the outcome favored the other party?
- (6) Were you comfortable with outcomes that favored the other party over you?

2.1.3. Participants. We ran the experiment with 81 participants recruited from Amazon Mechanical Turk, which is a crowdsourcing platform that allows people to complete online tasks in exchange for pay. Previous research shows that studies performed on Mechanical Turk can yield high-quality data, minimize experimental biases, and successfully replicate the results of behavioral studies performed on traditional pools [Paolacci et al. 2010]. We only sampled participants from the United

¹ Using deception allowed us to have precise experimental control over the counterpart’s behavior. Other researchers have also used deception when comparing behavior with humans vs. computers [Gallagher et al. 2002], [Rilling et al. 2002], [Sanfey et al. 2003], [Kircher et al. 2009].

States with an excellent performance history (95% approval rate in previous Mechanical Turk's tasks). Regarding gender, 28.4% of the participants were males. Age distribution was as follows: 18 to 21 years, 1.2%; 22 to 34 years, 55.6%; 35 to 44 years, 22.2%; 45 to 54 years, 9.9%; 55 to 64 years, 8.6%; over 65 years, 2.5%. Professional backgrounds were quite diverse. Participants were paid a fixed \$1.00 participation fee and also had the chance to win extra money through the lottery. Average time to complete the task was approximately 20 minutes. Participants gave their consent before engaging in the experiment. The research presented here was approved by the Internal Review Board at USC (ID# UP-14-00177).

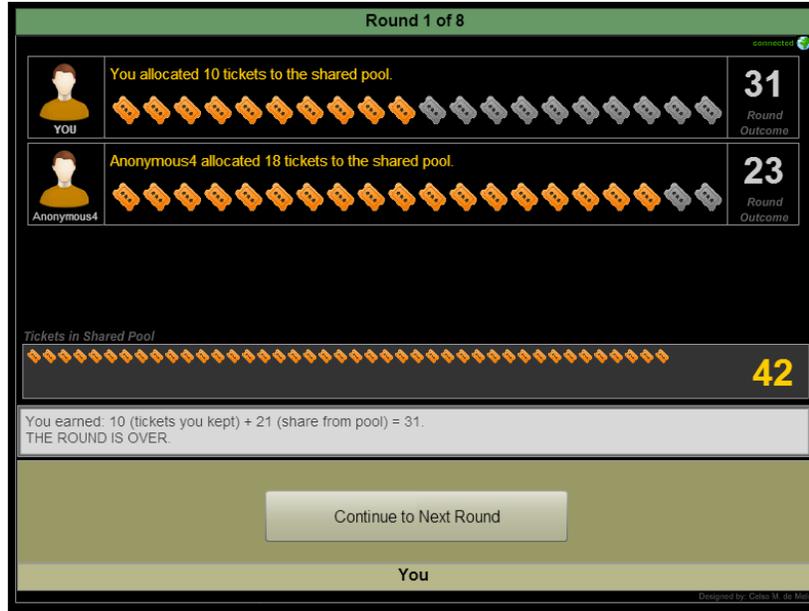


Fig. 1. The software used in Experiment 1.

2.2 Results

The means and standard errors for all dependent variables are shown in Table I. To compare average offer to human and computer counterparts, we ran an independent t test. This analysis revealed that participants allocated more money to the shared pool with humans than computers, $t(79) = 2.316$, $p = .023$, $r = .252$.

To analyze the subjective measures, we first ran a principal component analysis (varimax rotation), which confirmed the existence of two factors: one for guilt and one for envy. Guilt explained 31.3% of the variance and had the following loadings (Cronbach's $\alpha = .636$): "Were you concerned with not taking advantage of the other party?", .832; "Did you feel guilty when the outcome favored you?", .753; and, "Were you comfortable with outcomes that favored you over the other party?", -.695. Envy explained 30.5% of the variance and had the following loadings (Cronbach's $\alpha = .638$): "Did you feel envious when the outcome favored the other party?", .832; "Were you comfortable with outcomes that favored the other party over you?", -.797; and, "Were you concerned with not being taken advantage of?", .639. To compare the experience of emotion with each counterpart, we then ran independent t tests on each of these factors. The results showed that people experienced more guilt with humans than computers, $t(79) = 3.000$, $p = .004$, $r = .320$; however, there was no statistically

significant difference in experience of envy between humans and computers, $t(79) = 1.293$, $p = .200$.

Table I. Means and Standard Deviations for Decisions and Experience of Emotion in Experiment 1

	<i>n</i>	Human		Computer	
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
<i>Decision Measures</i>					
Allocation to Shared Pool in Public Goods Game	81	11.89	5.78	8.85	6.05
<i>Experience of Emotion Measures</i>					
Guilt Factor	81	.31	.85	-.33	1.06
Envy Factor	81	-.14	1.05	.15	.93

2.3 Discussion

The results show that people can treat machines in a social manner [Nass and Moon 2000], [Nass et al. 1996, 1997, 1999, 2000], [Reeves and Nass 1996]. Effectively, people's decisions with machines were, in general, more favorable than what selfish models of rational behavior predict. Nevertheless, in line with our prediction, the results also show that people favored humans to machines in terms of their offers to the shared pool. This confirms that people act differently when engaging in social interaction with machines. Also in line with our predictions, participants showed more guilt when exploiting humans, when compared to machines; however, contrary to our initial expectation, the results revealed that participants experienced just as much envy with humans as with machines.

3. EXPERIMENT 2

In this experiment, we wanted to replicate the findings in Experiment 1 in two novel tasks: the ultimatum and dictator games. Additionally, instead of using subjective measures, we used behavioral measures for the experience of guilt and envy. To accomplish this, we followed the methodology of Blanco et al. [2011]. In their method, participants engaged, in a repeated measures design, in the ultimatum game (as proposers and responders) and in a modified version of the dictator game.

In the modified dictator game, participants have to decide between keeping all of their initial endowment (20 lottery tickets) or accepting an even-split option. In our case, participants had to make twenty-one sequential decisions; each decision consisted of a choice between the option where they keep all 20 tickets – i.e., (20, 0) – or an even-split option – (0, 0), (1, 1)... (20, 20). Figure 2 shows a visual description of the game. For instance, the third decision in the sequence consisted of choosing between an option where the participant keeps all the 20 tickets vs. an option where each player gets 2 tickets. In this case, most participants in our experiment chose to keep all the tickets for themselves. In fact, most participants only started choosing the option to split when this gave each player at least 15 tickets. Moreover, once participants started splitting the tickets, they usually continued to split in all subsequent decisions. Notice that participants with monotone preferences should only have a single switching point (if any) where they change from choosing the left (20, 0) option to choosing the right even-split option.

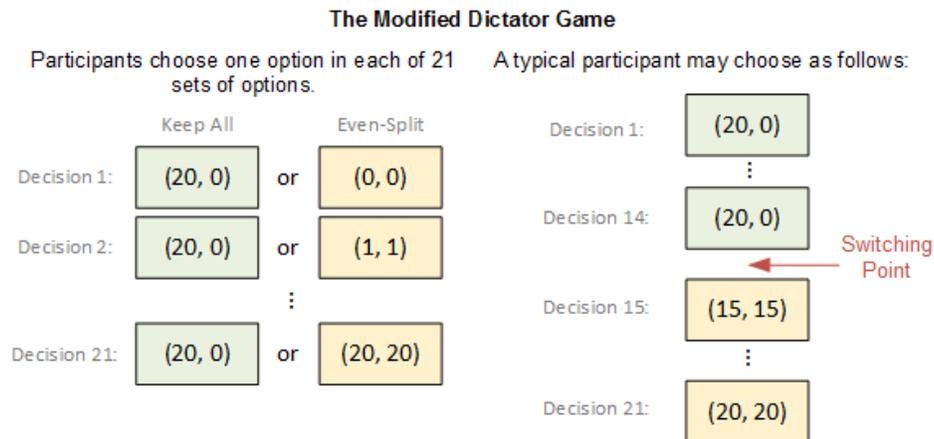


Fig. 2. The modified dictator game.

After completing these tasks, the data was fitted to Fehr and Schmidt's [1999] inequity aversion model. As described Section 1.3, this model summarizes participants' behavior according to two parameters that capture distinct forms of inequity aversion: one that represents aversion to disadvantageous inequality or envy (the α parameter), and one that represents aversion to advantageous inequality or guilt (the β parameter). The experimental data, first, allowed us to estimate these two parameters; subsequently, in line with the approach followed by other researchers (e.g., [Krajbich et al. 2009]), these parameters were interpreted as being behavioral measures of envy and guilt and, thus, were used to compare people's emotion experience when engaging with human vs. computer counterparts.

3.1 Method

3.1.1. Design. The experiment followed a 2×2 repeated-measures factorial design: *Counterpart* (Human vs. Computer) \times *Game* (Ultimatum Game vs. Modified Dictator Game). The counterpart and game order was counterbalanced across participants. Regarding the ultimatum game, participants played in both the role of proposer and responder (in a counterbalanced order). The initial endowment was 20 lottery tickets. The lottery tickets had financial consequences as they would enter a lottery for \$50; thus, the more tickets participants got, the higher were their chances of winning the lottery. Moreover, participants were informed that computers would enter the lottery. Participants were quizzed on these instructions and also engaged in an interactive tutorial, prior to starting the game.

Following the methodology of Blanco et al. [2011], we used the strategy-elicitation method. In this method, which has been extensively used in the past [Brandts and Charness 2011], participants report a choice for each possible decision node in the game. After reporting all choices, participants are randomly matched with a counterpart and assigned a role. The game is then simulated according to the reported decisions. This method contrasts with the direct-elicitation method, where participants are matched with other participants and engage in the game in real time. The advantage of the strategy-elicitation method is that it gathers more information per participant. This was also critical in our case, as decisions were necessary for all choice nodes in order to calculate the parameters in Fehr and Schmidt's model. Using this method also affected the logic of the lottery. Specifically, participants were instructed that, once all decisions were reported, the lottery would

proceed as follows: (a) Participants are assigned a random role; (b) The counterpart is determined randomly to be a human or a computer; (c) The game is simulated according to the reported decisions and, if the counterpart is a computer, an algorithm that makes decisions “just like real people” is used; (d) Tickets are assigned to the players according to the outcome of the simulated game; and (e) A lottery winner is selected, according to the odds set by the lottery tickets. Figure 3 and 4 show screenshots of the software.

a, Ultimatum Game (Proposer)

If you were the Sender and the Receiver was a computer agent



Sender: **YOU** Receiver: 
Computer Agent

how many tickets would you offer?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

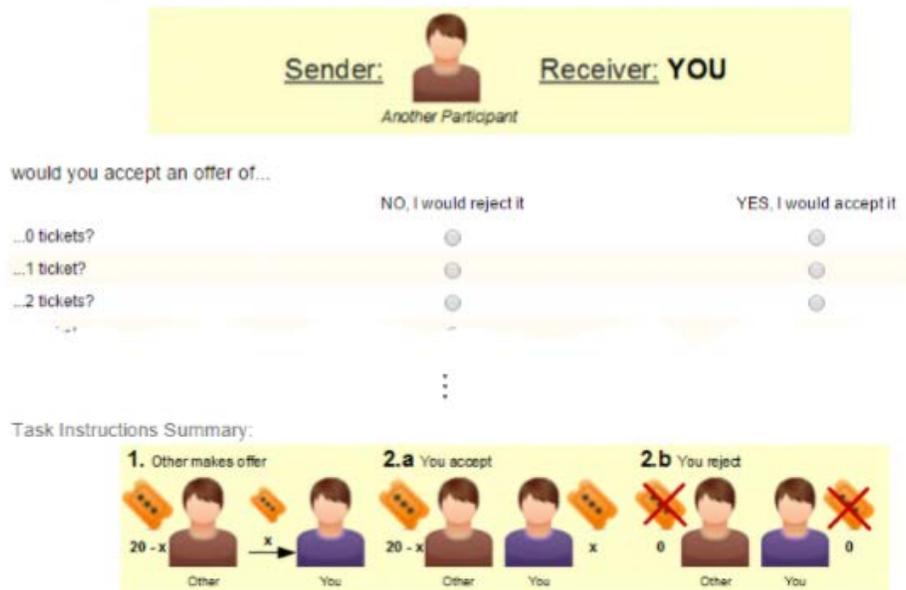
Task Instructions Summary:

1. You make offer **2.a Other accepts** **2.b Other rejects**

b, Ultimatum Game (Responder)

If you were the Receiver and the Sender was another participant



Sender:  Receiver: **YOU**
Another Participant

would you accept an offer of...

...0 tickets? NO, I would reject it YES, I would accept it

...1 ticket?

...2 tickets?

Task Instructions Summary:

1. Other makes offer **2.a You accept** **2.b You reject**

Fig. 3. The software used in Experiment 2 for the ultimatum game.

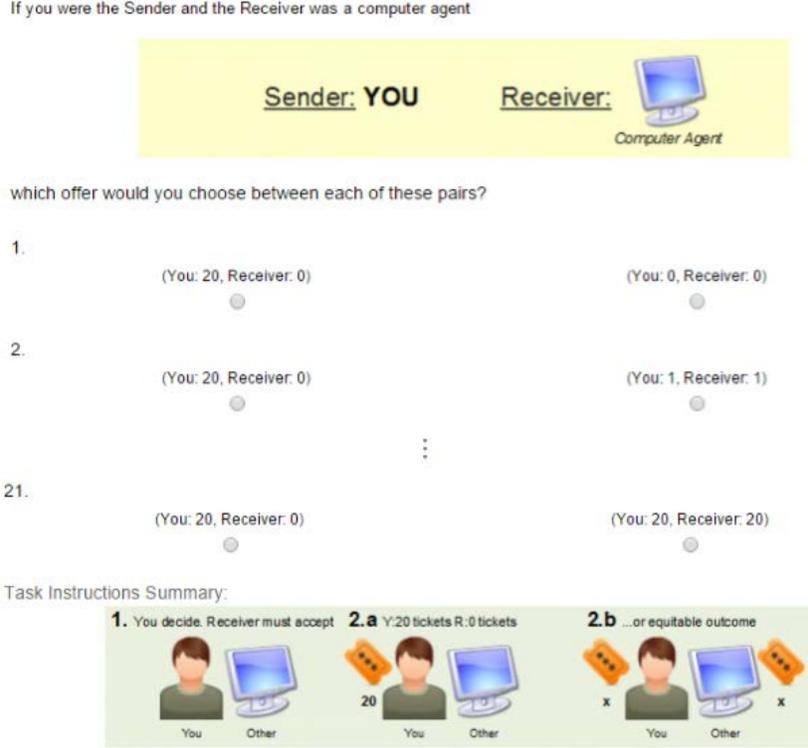


Fig. 4. The software used in Experiment 2 for the modified dictator game.

3.1.2. Dependent Variables. We focused on the following dependent variables: (a) The proposer's offer in the ultimatum game; (b) For responders in the ultimatum game, the point at which participants switched from rejecting to accepting the proposer's offers; (c) In the modified dictator game, the point at which participants switched from choosing (20, 0) to the split option.

To calculate the behavioral measures for envy (α parameter) and guilt (β parameter), we used the procedure from Blanco et al. [2011]. To calculate the α parameter, we used the data for responders in the ultimatum game. Suppose r'_k is the lowest offer responder k is willing to accept and, thus, $r'_k - 1$ is the highest offer rejected. A responder will hence be indifferent between accepting some offer $r_k \in [r'_k - 1, r'_k]$ and getting a zero payoff for rejection. This means that $U_i(r_k, 20 - r_k) = r_k - \alpha_i(20 - r_k - r_k) = 0$. Thus, the estimate for the envy parameter is:

$$\alpha_k = \frac{r_k}{2(10 - r_k)} \quad (2)$$

where we set $r_k = r'_k - 0.5$ as the point at which responders switched from rejecting to accepting offers. Participants that had more than one switch point – i.e., their preferences were inconsistent – were excluded from analysis.

The β parameter is calculated based on the decisions in the modified dictator game. We get β_k by finding the egalitarian allocation, (x_k, x_k) , such that the dictator is indifferent between keeping the entire endowment, the (20, 20) outcome, and (x_k, x_k) . Suppose that an individual switches to the egalitarian outcome at (x'_k, x'_k) ; i.e., s/he prefers (20, 0) over $(x'_k - 1, x'_k - 1)$ but (x'_k, x'_k) over (20, 0). Therefore, s/he is indifferent between the (20, 0) distribution and $(\tilde{x}_k, \tilde{x}_k)$, where $\tilde{x}_k \in [x'_k - 1, x'_k]$ and $x'_k \in \{1, \dots, 20\}$. Thus, $U_i(20, 0) = U_i(\tilde{x}_k, \tilde{x}_k)$ if, and only if, $20 - 20\beta_k = \tilde{x}_k$. This leads to:

$$\beta_k = 1 - \frac{\tilde{x}_k}{20} \quad (3)$$

where we set $\tilde{x}_k = x'_k - 0.5$ as the point at which people switched from choosing (20, 0) to choosing the even-split option. Participants that had more than one switch point – i.e., their preferences were inconsistent – were excluded from analysis. Finally, for participants that always chose (20, 0) – i.e., never split – we set \tilde{x}_k to 21².

3.1.3. Participants. Participants were recruited at the Psychology student pool of the University of Southern California. We recruited a total of 165 participants. We ran 6 participants at a time, throughout several sessions. Regarding gender, 53.0% were males. Age distribution was as follows: 21 years and under, 83.0%; 22 to 34 years, 17.0%. All participants were undergraduate or graduate students majoring in Psychology-related courses and mostly with citizenship from the United States (78.8%). Regarding incentive, 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 tickets as possible, as the total amount of tickets would increase their chances of winning the lottery. All participants gave informed consent before starting the task. The research presented here was approved by the Internal Review Board at USC (ID# UP-12-00350).

3.2 Results

To compare people's decisions with human and computer counterparts, we ran repeated-measures *t* tests. To compare the experience of guilt and envy, we used formulas (2) and (3) to calculate the α and β parameters with human and computer counterparts, and then ran repeated measures *t* tests. The means and standard errors for all these dependent variables are shown in Table II.

Table II. Means and Standard Deviations for Decisions and Experience of Emotion in Experiment 2

	<i>n</i>	Human		Computer	
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
<i>Decision Measures</i>					
Offers (UG, proposer)	165	7.91	3.05	6.71	3.43
Minimum accepted offer (UG, responder)	157 ^a	7.36	3.18	7.62	3.76
Switching point from (20, 0) to even-split (MDG)	143 ^b	15.60	6.87	18.13	5.86
<i>Experience of Emotion Measures</i>					
Guilt	140 ^{a,b}	1.77	4.85	1.98	4.83
Envy	140 ^{a,b}	.24	.34	.11	.29

UG – Ultimatum Game, MDG – Modified Dictator Game

^a Participants that switched from rejecting to accepting at more than one point (i.e., were inconsistent) were excluded.

^b Participants that switched from keeping all to the even-split at more than one point (i.e., were inconsistent) were excluded.

3.2.1. Ultimatum Game (Proposer). We compared participants' offers to humans and computers in the ultimatum game. The results showed that people offered more to humans than computers, $t(164) = 5.538$, $p = .000$, $r = .397$, mean difference = 1.20, 95% CI [.77, 1.63].

3.2.2. Ultimatum Game (Responder). We compared the point at which participants switched, with computers and humans, from rejecting to accepting offers in the

² Notice that this means that $\beta < 0$, which is an extension of Fehr and Schmidt's assumptions. People that fall on this category may enjoy being better off than others.

ultimatum game. Participants that had more than one switching point – i.e., had inconsistent preferences – were excluded. We excluded 8 participants using this criterion. The results showed that participants did not show a statistically significant difference in switching point between humans and computers, $t(156) = 1.184$, $p = .238$.

3.2.3. Modified Dictator Game. We compared the point at which participants switched, with humans and computers, from choosing (20, 0) to the even-split in the modified dictator game. Participants that had more than one switching point – i.e., had inconsistent preferences – were excluded. We excluded 22 participants using this criterion. Participants that never switched to the even split – i.e., always chose (20, 0) – were assigned a switch point of 21 for this analysis. The results showed that participants switched to the even-split later with computers than with humans, $t(142) = 5.590$, $p = .000$, $r = .425$, mean difference = 2.52, 95% CI [1.63, 3.42].

3.2.4. Envy and Guilt. The results showed that people did not experience more envy (α parameter) with humans than with computers, $t(139) = .517$, $p = .606$; in contrast, people experienced more guilt (β parameter) with humans than with computers, $t(139) = 5.603$, $p = .000$, $r = .429$, mean difference = .13, 95% CI [.08, .17].

3.3 Discussion

The results in Experiment 2 reinforced the findings from Experiment 1. People were more generous in their offers in the ultimatum game to humans than to computers. Moreover, in the modified dictator game, people required more tickets in the even split with computers than with humans, before they were willing to forfeit keeping all tickets. The only case where participants did not seem to distinguish between humans and computers was when they assumed the role of responders in the ultimatum game, in which case they accepted offers at the same level. We can gather insight on these results by looking at the behavioral measures of envy and guilt. As in the previous experiment, people showed more guilt with humans than with computers, but did not show a difference in experience of envy. This might explain why participants were willing to exploit computers in the ultimatum game (as proposers) and in the modified dictator game, while at the same time being willing to accept the same level of offer in the ultimatum game (as responders) from humans and computers.

4. GENERAL DISCUSSION

In line with work by Nass and colleagues [Nass and Moon 2000], [Nass et al. 1996, 1997, 1999, 2000], [Reeves and Nass 1996], this paper confirms that people show social consideration for machines, when immersed in social interaction. Contrary to predictions by standard rational models of decision making, even when engaging with machines, people allocated money to a shared pool in the public goods game and made non-zero offers in the ultimatum and dictator games. Nevertheless, the paper reinforces that people still make important distinctions between machines and humans. Everything else being equal, people made more favorable decisions to humans than to machines: (1) people allocated more money to a shared public good with humans, (2) people offered more money to humans in the ultimatum game, and, (3) in the modified dictator game, people expected more money from machines than from humans, before they were willing to forfeit an option where they kept everything vs. an option that gave both the same number of tickets.

We presented consistent evidence for this behavioral difference in three economic games – public goods, ultimatum, and dictator games – and with participants from a general online sample and a student sample. Nevertheless, some may still argue that this difference occurred because people were avoiding giving lottery tickets to machines, since if they won, no one else could win it. However, we can present three reasons why this explanation is unlikely: (a) if people truly wanted to minimize the chances of machines winning the lottery, then they should have behaved according to the predictions of selfish models (e.g., offered 1 ticket as proposers in the ultimatum game); (b) people should have also avoided giving tickets to humans, as they can also win the lottery; (c) several studies [Gallagher et al. 2002], [Kircher et al. 2009], [Krach et al. 2008], [McCabe et al. 2001], [Nass and Moon 2000], [Nass et al. 1996, 1997, 1999, 2000], [Reeves and Nass 1996], [Rilling et al. 2002], [Sanfey et al. 2003], [van't Wout et al. 2006] demonstrate that people show social considerations for machines and, in many cases, this happens automatically.

Another concern may have been that machines were interpreted as proxies for the experimenters or the programmers. In this case, the reported effects would have been due to the distinction between interacting with an individual vs. a group (e.g., people tend to act more selfishly in impersonal market environments [Roth et al. 1991]), or due to a perceived difference in group membership – i.e., people perceived participants as being in-group members, whereas experimenters were out-group members. We advance two reasons why this alternative seems implausible: (a) Earlier research shows that people are able to make clear distinctions between interacting with computers and the people responsible for programming them [Nass et al. 1994; Study 5]; and, (b) The instructions described the computers as acting on their own behalf and never mentioned the experimenters.

To gain insight into these effects, we looked at people's experience of guilt and envy. Experiment 1 showed that people self-reported experiencing more guilt with humans than with machines, when the economic outcome was advantageous in their favor; in contrast, people self-reported experiencing as much envy with machines as with humans, when the economic outcome was disadvantageous to them. These results were further replicated in Experiment 2 using behavioral measures of guilt and envy. To the best of our knowledge, this is the first paper to show the role of guilt (and the absence of a role of envy) in explaining differences in behavior with machines, when compared to humans.

The fact that people were more likely to experience guilt with humans than with machines helps explain why they made more favorable offers with humans. Previous research shows that people try to avoid harming others to prevent the unpleasant experience of guilt [Carlsmith and Gross 1969], [Freedman et al. 1967], [Ketelaar and Au 2003] [Van Kleef et al. 2006]. This mechanism, thus, introduces an emotional cost to acting selfishly. However, if people experience less guilt when engaging with machines, then they will consequently be more comfortable with exploiting them. But, why do people experience less guilt with machines? One possibility follows from research showing that people expect machines to have lower mental ability than humans, in particular with respect to the ability to sense and experience emotion [Gray et al. 2007], [Waytz et al. 2010]. Future work should, thus, test whether this difference in experience of guilt will disappear when engaging with machines that are capable of recognizing, synthesizing, and expressing emotion [Marsella et al. 2010], [Picard 2010].

The absence of a difference in experience of envy, in contrast, suggests that people do not like being exploited, independently of whether others are humans or not. The

critical difference, thus, was that the target of the unfairness for envy is the self, whereas for guilt the target is the other. Moreover, contrary to prior research [Lim and Reeves 2010], [Mandryk et al. 2006], [Okita et al. 2007], [Ravaja 2009], [Sanfey et al. 2003], [Van't Wout et al. 2006], this result emphasizes that people do not always experience less emotion with machines than with humans.

This paper relied on Fehr and Schmidt's [1999] model to gather insight on people's experiences in human-machine interaction. Nevertheless, experimental economists have proposed several other interesting models [Bolton and Ockenfels 2000], [Levine 1998], [Rabin 1993], [Falk and Fischbacher 2006]. Fehr and Schmidt fall within the class of inequity aversion models. Another class of models that is likely to be useful is one that focuses on the intentions underlying (fair or unfair) outcomes. These models argue that what is critical is not the outcome itself, but the intentions that led to it. For instance, a counterpart might have chosen a rather unfair outcome, but if s/he did not have a reasonable alternative, people are much more likely to accept the outcome [Rabin 1993]. Falk and Fischbacher [2006], in turn, attempt to integrate inequity aversion with intentions and propose a model that is sensitive both to the outcome as well as to the underlying intentions. A potentially interesting line of inquiry, thus, could explore whether people's behavior and emotional experiences are influenced by perceived intentions in machines, in the same manner as in humans. Overall, we propose that these economic models have much to offer to human-machine interaction research, in particular, for the study of the mechanisms underlying people's decision making with machines, when compared to humans.

Finally, the results presented in this paper have important implications for the design of human-machine interaction systems. Designers need to be aware that people behave differently, at least by default, with machines when compared to humans. Thus, if the intention is to minimize this difference, designers can take appropriate action. First, they can try to de-emphasize that people are interacting with an autonomous machine; however, there may be ethical and legal concerns that limit this type of approach. For instance, the UK's 1998 Data Protection Act gives employees the right to ask for human intervention in the case of any decision made solely by automated means, when personal data is involved. A better alternative, thus, may be to emphasize that machines are acting on behalf of real humans. Yet another alternative may be to increase the anthropomorphic cues in machines. For instance, one way to make people experience more emotion with machines is to simulate expression of emotion [Ahyoung et al. 2015], [de Melo et al. 2012, 2014]. Previous work shows that emotion in machines can have a powerful positive effect on people's behavior [Marsella et al. 2010], [Picard 1997]. Lastly, the findings reported here are not relevant exclusively to software agents but also to other kinds of machines, such as social robots [Breazeal 2003]. Research shows that, just like for computers, people have lower expectations about robots' mental capacities, when compared to humans [Gray et al. 2007], [Waytz et al. 2010]. Thus, we expect people to show similar differences in behavior with robots and, thus, if the goal is to minimize such differences, appropriate measures should be taken. Overall, this paper shows that we cannot afford to ignore emotional factors – especially, the experience of guilt – when trying to make sense of people's behavior with machines; but, if we take these considerations into account when designing human-machine interaction systems, then we can hope to achieve the same level of efficiency – cooperation, fairness, reciprocity, altruism, etc. – we see in human-human interaction.

REFERENCES

- R. Axelrod, and W. Hamilton. 1981. The evolution of cooperation. *Science* 211, 1390-1396.
- M. Bertrand, and S. Mullainathan. 2001. Do people mean what they say? Implications for subjective survey data. *The American Economic Review* 91, 67-72.
- M. Blanco, D. Engelmann, and H. Normann. 2011. A within-subject analysis of other-regarding preferences. *Games and Economic Behavior* 72, 321-338.
- J. Blascovich, J. Loomis, A. Beall, K. Swinth, C. Hoyt, and J. Bailenson. 2002. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry* 13, 103-124.
- G. Bolton and A. Ockenfels. 2000. A theory of equity, reciprocity and competition. *The American Economic Review* 100, 166-193.
- J. Brandts and G. Charness. 2011. The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics* 14, 375-398.
- C. Breazeal. 2003. Toward sociable robots. *Robotic and Autonomous Systems* 42, 167-175.
- J. Carlsmith, and A. Gross. 1969. Some effects of guilt on compliance. *Journal of Personality and Social Psychology* 11, 232-239.
- A. Choi, C. de Melo, P. Khooshabeh, W. Woontack, and J. Gratch. 2015. Physiological evidence for a dual process model of the social effects of emotion in computers. *International Journal of Human-Computer Studies* 74, 41-53.
- A. Damasio. *Descartes' error: Emotion, reason, and the human brain*. Putnam, New York, NY.
- T. Davenport, and J. Harris. 2005. Automated decision making comes of age. *Sloan Management Review* 46, 83-89.
- C. de Melo, P. Carnevale, and J. Gratch. 2011. The effect of expression of anger and happiness in computer agents on negotiations with humans. In *Proceedings of Autonomous Agents and Multiagent Systems (AAMAS'11)*. 937-944.
- C. de Melo, P. Carnevale, and J. Gratch. 2012. The impact of emotion displays in embodied agents on emergence of cooperation with people. *Presence* 20, 449-465.
- C. de Melo, P. Carnevale, S. Read, and J. Gratch. 2014. Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology* 106, 73-88.
- R. Dawes. 1980. Social dilemmas. *Annual Review of Psychology* 31, 169-193.
- M. Duffy, K. Scott, J. Shaw, B. Tepper, and K. Aquino. 2012. A social context model of envy and social undermining. *Academy of Management Journal* 55, 643-666.
- A. Falk and U. Fischbacher. 2006. A theory of reciprocity. *Games and Economic Behavior* 54, 293-315.
- E. Fehr, and S. Gächter. 2000. Cooperation and punishment in public goods experiments. *The American Economic Review* 90, 980-994.
- E. Fehr, and K. Schmidt. 1999. *Quarterly Journal of Economics* 114, 817-868.
- E. Fehr, and K. Schmidt. 2006. The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. In S-C. Kolm, and J. Ythier (Eds.). *Handbook of the Economics of Giving, Altruism and Reciprocity*. Elsevier, Oxford, UK, 615-691.
- R. Forsythe, J. Horowitz, N. Savin, and M. Sefton. 1994. Fairness in simple bargaining games. *Games and Economic Behavior* 6, 347-369.
- J. Freedman, S. Wallington, and E. Bless. 1967. Compliance without pressure: The effect of guilt. *Journal of Personality and Social Psychology* 7, 117-124.
- H. Gallagher, J. Anthony, A. Roepstorff, and C. Frith. 2002. Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814-821.
- J. Gratch, N. Wang, J. Gerten, E. Fast, and R. Duffy. 2007. Creating rapport with virtual agents. In *Proceedings of Intelligent Virtual Agents (IVA'07)*. Springer Berlin Heidelberg, 125-138.
- H. Gray, K. Gray, and D. Wegner. 2007. Dimensions of mind perception. *Science* 315, 619.
- W. Güth, R. Schmittberger, and B. Schwarze. 1982. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* 3, 367-388.
- N. Haslam. Dehumanization: An integrative review. 2006. *Personality and Social Psychology Review* 10, 252-264.
- J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. *The American Economic Review* 91, 73-78.
- E. Hoffman, K. McCabe, K. Shachat, and V. Smith. 1994. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior* 7, 346-380.
- M. Isaac, and J. Walker. 1988. Group size effects in public goods provision: The voluntary contribution mechanism. *Quarterly Journal of Economics* CIII, 179-199.
- S. Jahedi, and F. Méndez. 2014. On the advantages and disadvantages of subjective measures. *Journal of Economic Behavior & Organization* 98, 97-114.
- N. Jennings, P. Faratin, A. Lomuscio, S. Parsons, M. Wooldridge, C. Sierra. 2001. Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation* 10, 199-215.

- T. Ketelaar, and W. Au. 2003. The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion* 17, 429-453.
- T. Kircher, I. Blümel, D. Marjoram, T. Lataster, L. Krabbendam, J. Weber, J. van Os, and S. Krach. 2009. Online mentalising investigated with functional MRI. *Neuroscience Letters* 454, 176-181.
- S. Krach, F. Hegel, B. Wrede, G. Sagerer, F. Binkofski, and T. Kircher. 2008. Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLOS ONE* 3, 1-11.
- I. Krajbich, R. Adolphs, D. Tranel, N. Denburg, and C. Camerer. 2009. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *The Journal of Neuroscience* 29, 2188-2192.
- P. Kollock. 1998. Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology* 24, 183-214.
- K. Leith, and R. Baumeister. 1998. Empathy, shame, guilt, and narratives of interpersonal conflicts: Guilt-prone people are better at perspective taking. *Journal of Personality and Social Psychology* 66, 1-37.
- J. Lerner, Y. Li, P. Valdesolo, and K. Kassam. 2015. Emotion and decision making. *Annual Review of Psychology* 66, 799-823.
- D. Levine. 1998. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1, 593-622.
- M. Lewis. 2000. Self-conscious emotions: Embarrassment, pride, shame, and guilt. In M. Lewis, and J. Haviland-Jones (Eds.). *Handbook of emotions*, 2nd edn. Guilford Press, New York, 623-636.
- S. Lim, and B. Reeves. 2010. Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human-Computer Studies* 68, 57-68.
- R. Lin, and S. Kraus. 2010. Can automated agents proficiently negotiate with humans? *Communications of the ACM* 53, 78-88.
- G. Loewenstein, and J. Lerner. 2003. The role of affect in decision making. In R. Davidson, K. Scherer, and H. Goldsmith (Eds.). *Handbook of Affective Sciences*. Oxford University Press, New York, NY, 619-642.
- R. Mandryk, K. Inkpen, and T. Calvert. 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology* 25, 141-158.
- S. Marsella, J. Gratch, and P. Petta. 2010. Computational models of emotion. In K. Scherer, T. Bänziger, and E. Roesch (Eds.). *A blueprint for an affectively competent agent: Cross-fertilization between Emotion Psychology, Affective Neuroscience, and Affective Computing*. Oxford University Press, New York, NY.
- K. McCabe, D. Houser, L. Ryan, V. Smith, and T. Trouard. 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Science* 98, 11832-11835.
- C. Nass, and Y. Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 56, 81-103.
- C. Nass, B. Fogg, and Y. Moon. 1996. Can computers be teammates? *International Journal of Human-Computer Studies* 45, 669-678.
- C. Nass, K. Isbister, and E.-J. Lee. 2000. Truth is beauty: Researching embodied conversational agents. In J. Cassell (Ed.). *Embodied conversational agents*. MIT Press, Cambridge, MA, 374-402.
- C. Nass, Y. Moon, and P. 1999. Carney. Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Psychology* 29, 1093-1110.
- C. Nass, Y. Moon, and N. Green. 1997. Are computers gender-neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology* 27, 864-876.
- C. Nass, J. Steuer and E. Tauber. 1994. Computers are social actors. In *Proceedings of Human Factors in Computing Systems (CHI94)*. 72-78.
- S. Okita, J. Bailenson, and D. Schwartz. 2008. Mere belief in social action improves complex learning. In *Proceedings of the 8th International Conference on for the Learning Sciences (ICLS'08)*. 132-139.
- G. Paolacci, J. Chandler, and P. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 411-419.
- R. Picard. 1997. *Affective computing*. The MIT Press, Cambridge, MA.
- M. Pillutla, and J. Murnighan. 1996. Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes* 68, 208-224.
- D. Pruitt, and M. Kimmel. 1977. Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology* 28, 363-392.
- M. Rabin. 1993. Incorporating fairness into game theory and economics. *The American Economic Review* 83, 1281-1302.
- N. Ravaja. 2009. The psychophysiology of digital gaming: The effect of a non co-located opponent. *Media Psychology* 12, 268-294.
- B. Reeves, and C. Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
- J. Rilling, D. Gutman, T. Zeh, G. Pagnoni, G. Berns, and C. Kilts. 2002. A neural basis for social cooperation. *Neuron* 35, 395-405.

- A. Roth, P. Vesna, O.-F. Masahiro and Z. Shmuel. 1991. Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review* LXXXI, 1068-1095.
- A. Sanfey, J. Rilling, J. Aronson, L. Nystrom, and J. Cohen. 2003. The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755-1758.
- R. Smith, and S. Kim. 2007. Comprehending envy. *Psychological Bulletin* 133, 46-64.
- J. Tangney. 1990. Assessing individual differences in proneness to shame and guilt: Development of the self-conscious affect and attribution inventory. *Journal of Personality and Social Psychology* 59, 102-111.
- J. Tangney. 1995. Shame and guilt in interpersonal relations. In J. Tangney, and K. Fischer (Eds.). *Self-conscious emotions: The psychology of shame, guilt, embarrassment, and pride*. Guilford Press, New York, 114-139.
- J. Tangney. 1999. The self-conscious emotions: Shame, guilt, embarrassment, and pride. In T. Dalgleish, and M. Power (Eds.). *Handbook of cognition and emotion*. Wiley, New York, 541-568.
- G. Van Kleef, C. De Dreu, and A. Manstead. 2006. Supplication and appeasement in conflict and negotiation: The interpersonal effects of disappointment, worry, guilt, and regret. *Journal of Personality and Social Psychology* 91, 124-142.
- M. van't Wout, R. Kahn, A. Sanfey, and A. Aleman. 2006. Affective state and decision-making in the ultimatum game. *Experimental Brain Research* 169, 564-568.
- A. Waytz, K. Gray, N. Epley, and D. Wegner. 2010. Causes and consequences of mind perception. *Trends in Cognitive Science* 14, 383-388.