Social Categorization and Cooperation between Humans and Computers

Celso M. de Melo (demelo@usc.edu)
Peter J. Carnevale (peter.carnevale@marshall.usc.edu)
USC Marshall School of Business
Los Angeles, CA 90089-0808 USA

Jonathan Gratch (gratch@ict.usc.edu)
USC Institute for Creative Technologies, 12015 Waterfront Drive, Building #4
Playa Vista, CA 90094-2536 USA

Abstract

Computers increasingly perform a variety of important tasks and services that influence individuals and organizations, yet few studies tell us about how humans interact with computers and other non-human decision-makers. In four experiments, we asked people to engage in cooperation tasks with computers and with humans. Experiment 1 found that people gave more money to a human than a computer. We argue this effect reflects a basic bias in favor of humans, which are perceived to be the in-group, when compared to computers, which are perceived to be the out-group. In Experiment 2, we varied computer and human ethnicity to be the same or different as the participant; results indicated that ethnicity had a parallel but additive effect that was independent to the effect of the human social category. The data of Experiment 3 indicate that it is also possible to promote group membership with computers by creating structural interdependence based on shared incentives. Finally, we demonstrate in Experiment 4 that our framework based on social categorization theory can predict situations where people will cooperate more with computers than with humans. We discuss implications for understanding people's decision making with human and non-human others.

Keywords: Human vs. Computers; Decision Making; Cooperation; Social Categorization; Group Membership.

Introduction

Computers are routinely involved in decisions that affect the lives of individuals and organizations (Davenport & Harris, 2005). The remarkable growth of business conducted online (U.S. Census Bureau, 2011) and the increasing amounts of time people spend interacting via social media (Honignam, 2012), suggests computers will play an even more pervasive role on people's social, economic, and political life. As a consequence of these changes, people are faced with having to make decisions with, not human but, artificial decision makers. However, despite the importance of the issue, there has been remarkably little research on the nature of people's decision making with such non-human counterparts.

We propose that social categorization theory is a useful framework for understanding how humans reach decisions with computers (Tajfel & Turner, 1986; Turner, Oakes, Haslam, & McGarty, 1994; Crisp & Hewstone, 2007). One proposition of the theory is that people categorize others into groups while associating, or self-identifying, more with some (the in-groups) than others (the out-groups). Because of this categorization, people will conform more to the values and norms of the group, and tend to favor the in-group to the out-group – a phenomenon referred to as in-group bias. One consequence of this bias is that people trust and cooperate more with in-group than out-group members (Sherif, Harvey, White, Hood, & Sherif, 1961).

We propose that people make a basic distinction between humans and computers in terms of membership in the "human social category". People are capable of dehumanizing and subsequently discriminating against others that are perceived to lack certain mental abilities (Haslam. 2006). These abilities are of two types (Gray, Gray, & Wegner, 2007; Loughnan & Haslam, 2007): agency, the capacity to act and plan; and, experience, the capacity to sense and feel. Our proposal rests on research that shows that people perceive computers to possess less of these mental abilities than humans (Blascovich et al., 2002; Gray et al., 2007; Waytz, Gray, Epley, & Wegner, 2010). Further research shows that people tend to show stronger activation with humans, when compared to computers, of brain regions associated with mentalizing (i.e., the inferring of others' mental states) and the experience of affect (Gallagher, Anthony, Roepstorff, & Frith, 2002; Krach et al., 2008; McCabe, Houser, Ryan, Smith, & Trouard, 2001; Rilling et al., 2002; Sanfey, Rilling, Aronson, Nystrom, & Cohen; 2003). We, therefore, posit that denial of membership in the human social category to computers is due to this perceived difference in mental abilities. To test the existence of a bias in favor of humans, we compared participants' money offers to computers with offers to humans, in prototypical decision making tasks (Experiments 1, 2, and 3).

Human identities, however, are complex and multifaceted. In many settings, more than one social category (e.g., gender, age, ethnicity) may be relevant and influence behavior (Crisp & Hewstone, 2007). On the one hand, context can prime one category to become more dominant (or salient) and effectively exclude the influence of the others (e.g., Shih, Pittinsky, & Ambady, 1999). On the other hand, social categories can be simultaneously salient and have an additive effect on people's behavior (Crisp & Hewstone, 2007).

Our second proposal, therefore, is that multiple social categorization also applies in human-computer interaction. In particular, we posit that the effects of the human social category can combine in additive fashion with the effects of other social categories. Earlier research has already shown that people can apply human stereotypes to computers (Nass & Moon, 2000): in one experiment, in line with gender stereotypes, people assigned more competence to computers with a female voice than a male voice on the topic of "love and relationships"; in another experiment, people perceived computers with a virtual face of the same ethnicity as being more trustworthy and giving better advice than a computer with a face of a different ethnicity. Aside from manipulation of characteristics of the computers, research has also shown that it is possible to create group membership with computers by manipulating characteristics of the situation. Nass, Fogg, & Moon (1996) showed that people can favor a computer that belongs to the team, as defined by interdependence in the task's payoffs, when compared to a non-team computer. Thus, to test our proposal, we present several experiments where people engaged with humans and computers but, we also introduced additional social categories based on manipulation of the characteristics of the computers, namely ethnicity (Experiments 2 and 4), and of the situation, namely interdependence through shared payoffs (Experiments 3 and 4). We test whether the effect of the human social category on participants' offers dominates or combines, in additive fashion, with the effects of the other categories.

Experiment 1

Experiment 1 tests whether people make a basic distinction, in a decision making context, that favors humans when compared to computers. Participants engaged in a simple task, the dictator game (Forsythe, Horowitz, Savin, & Sefton, 1994), with computers that were perceived to be controlled by computer algorithms – agents – or computers that were perceived to be controlled by other participants – avatars. The dictator game involves two players: a sender and a receiver. The sender gets 20 tickets and decides how many to give the receiver, who has no choice but to accept it. The tickets had financial consequences, as they would go into a lottery for a prize in real money. Participants were told that agents would participate in the lottery and, if an agent won, no one would get the prize. Because there is no material incentive to offer anything, the game is seen as an index of altruism. Previous experiments show that 60% of participants tend to offer something and, on average, 28% of the pie is offered (Engel, 2011). In our experiment, participants always played the role of the sender and engaged, in a repeated measures design, with agents and avatars. We recruited 47 participants on Amazon Mechanical Turk for this experiment. To support the deception pertaining to avatars, before starting the task, participants were asked to wait for a "server to connect them to other participants"; moreover, while they waited they saw information that several other participants were already connected to the server. In fact, there was no server. Finally, in all our experiments, participants were supposedly matched with counterparts of the same gender. Participants were debriefed at the end regarding these deceptions.

To facilitate interpretation, we converted participants' offers into percentages (over the total amount of 20 tickets). As predicted, the results revealed that participants offered more tickets to avatars (M = 32.61, SD = 18.21) than to agents (M = 15.43, SD = 15.52), $t(46)$ = 6.702, p = .000, r =.703, mean difference = 17.18, 95% CI [12.02, 22.34].

The results confirm that people show a bias that favors humans to computers in the dictator game. We argue this occurs because computers are perceived to lack in certain mental abilities (Blascovich et al., 2002; Gray et al., 2007; Waytz et al., 2010) and, thus, are perceived as out-group members (in the human social category).

Experiment 2

In Experiment 2, we introduced a new social category – ethnicity of the counterpart – and tested whether its effect dominated, combined, or was dominated by the effect of the human social category. Although racial discrimination is on the decline (e.g., Ford, 2008), people tend to make automatic distinctions based on race, which can produce subtle forms of racial discrimination (Gaertner & Dovidio, 2005). In human-computer interaction, previous studies had also shown that computers with a visual representation corresponding to the same ethnicity were perceived more favorably (Nass & Moon, 2000; Rossen, Johnsen, Deladisma, Lind, & Lok, 2008). However, this earlier work focused on subjective impressions, whereas our experiment is the first, to the best of our knowledge, to test the effect of ethnicity on behavioral measures in a decision task.

Participants engaged in the dictator game, in a between-participants factorial design, with (perceived) agents or avatars that had a virtual face corresponding to either the same or different ethnicity as the participant. For instance, if the participant was Caucasian then, in the 'different ethnicity' condition, the counterpart's virtual face would be randomly chosen from one of the following: African-American, Hispanic, Southeast Asian, or East Indian. We tested three competing hypotheses: 1) if the human category is more salient than ethnicity, then people should not distinguish between agents of the same or different ethnicity; 2) if the ethnicity category is more salient than the human category, then people should not distinguish between agents and avatars of the same ethnicity; 3) if the human and ethnicity categories are independent, then they should lead to an additive effect on people's offers. Figure 1 shows the ethnicities we considered and some of the corresponding virtual faces. We recruited 184 participants at the USC Marshall School of Business. Most participants were Caucasian (32.1%) or Southeast Asian (56.0%). To support the deception related to avatars, we ran 12 participants per session, and instructed them that "other participants will be controlling" the avatars. Moreover, we also told them that they "would connect to a server that matches players with each other". Participants were debriefed at the end regarding this deception.

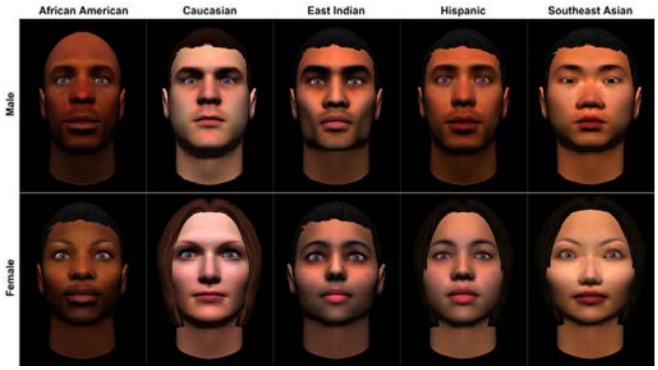

Figure 1: The ethnicities and some of the virtual faces used in Experiments 2 and 4.

To facilitate interpretation, we converted offers into percentages (over the total amount of 20 tickets). The average offers are shown in Figure 2. To analyze the data we ran an Other (Agent vs. Avatar) × Ethnicity (Same vs. Different) ANOVA. The results replicated the main effect of Other reported in Experiment 1: people offered more tickets to avatars (M = 20.63, SD = 20.27) than to agents (M = 15.44, SD = 17.37), $F(1, 180) = 4.08$, $p = .045$, partial $\eta^2 = .022$. The results also confirmed a main effect of Ethnicity: people offered more to counterparts of the same ethnicity (M = 23.26, SD = 20.21) than counterparts of a different ethnicity (M = 12.80, SD = 16.20), $F(1, 180) = 15.55$, $p = .000$, partial $\eta^2 = .080$. Finally, the results revealed no Other × Ethnicity interaction, $F(1, 180) = .347$, $p = .556$.

Thus, the results show that multiple social categories can be applied to computers; moreover, the results suggest that the human and ethnicity categories can have an independent and additive effect on participants' money offers.

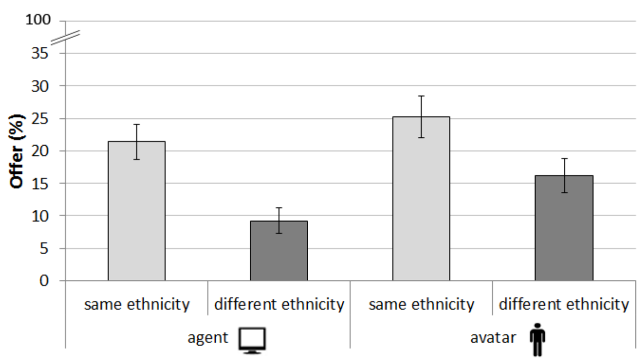

Figure 2: Average offers in Experiment 2. The error bars show the standard errors.

Experiment 3

In contrast to Experiment 2, which created group membership by manipulating a (visual) characteristic of the counterpart, Experiment 3 manipulated group membership by creating payoff interdependence among players. To achieve this we used a decision making task, the nested social dilemma (Wit & Kerr, 2002), which splits players into groups and bids group interests against collective interests. Generally, a social dilemma is a situation 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 (Dawes, 1980). Specifically, the nested social dilemma is a 6-player task where the participant is randomly allocated to position A, B, C, D, E or F and accordingly assigned to group ABC or DEF. The participant is given 30 lottery tickets that can be invested in three accounts: the *private*, *in-group* and *all* accounts. Tickets invested to the private account are multiplied by 1.0 and returned to the participant; tickets invested to the in-group account, which is referred to in the instructions as the "group account", are multiplied by 2.5 and split equally among all group members; tickets invested to the all account are multiplied by 4.0 and split equally by all six players. These payoff characteristics create interdependence among group members and preserve the defining properties of a social dilemma: irrespective of others' allocations, shifting points from a higher to a lower level account always increased one's individual final payoff; however, if everyone is selfish and invests in a lower account, then everyone is worse off than if they had invested in a higher account.

The experiment followed a 2 × 2 between-participants factorial design: In-Group (Agents vs. Avatars) × Out-Group (Agents vs. Avatars). In line with earlier work on the in-group bias (Sherif et al., 1961), we expected people to favor in-group avatars to out-group avatars. Since a previous study had already shown that people can favor a computer that belongs to the team when compared to a non-team computer (Nass et al., 1996), we also expected people to favor in-group agents to out-group agents. When engaging with in-group avatars and out-group agents, we expected people to strongly favor the in-group not only because they belonged to the interdependent group but also to the human social category. The last case is more interesting: when engaging with in-group agents and out-group avatars, interdependence favors the agents but people also identify with the human social category of the out-group. Following the results in the previous experiment we expected these two influences to cancel each other out, which would result in no preference between the in- and out-groups. We recruited 116 participants at the USC Marshall School of Business. To support the deception pertaining to avatars, we followed a similar procedure to Experiment 2.

To facilitate interpretation, we converted ticket allocations into percentages. As expected people invested more in the private than the other accounts (private: M = 66.23, SD = 26.38; in-group: M = 21.26, SD = 19.60; all: M = 12.51, SD = 16.92); however, to test our hypotheses we focused on a measure for the in-group bias, which we operationalize as the difference between allocations to the in-group and the all accounts. Figure 3 shows the means and confidence intervals for this measure for each condition. We then ran an In-Group × Out-Group ANOVA on this measure. The results revealed a statistically significant In-Group × Out-Group interaction, $F(1, 118) = 4.13$, $p = .044$, partial η^2 = .034. To further understand this interaction we ran one-way *t*-tests, for each condition, to test if in-group bias was different than zero. For in-group avatars and out-group avatars, in-group bias was statistically significantly different than zero, $t(28) = 2.63$, $r = .445$, mean difference = 12.64, 95% CI [2.79, 22.50]. For in-group agents and out-group agents, in-group bias was also statistically significant, $t(29) = 2.644$, $r = .441$, mean difference = 14.33, 95% CI [3.25, 25.42]. For in-group avatars and out-group agents, in-group bias was once more statistically significant, $t(31)$ = 3.73, $r = .557$, mean difference = 7.81, 95% CI [3.54, 12.08]. However, for in-group agents and out-group avatars, in-group bias was not statistically significantly different from zero, $t(30) = .122$, $r = .022$, mean difference = .65, 95% CI [-10.15, 11.44].

The results, thus, confirmed our prediction that people would favor the in-group in all cases, except when the in-group was composed of agents and the out-group of avatars.

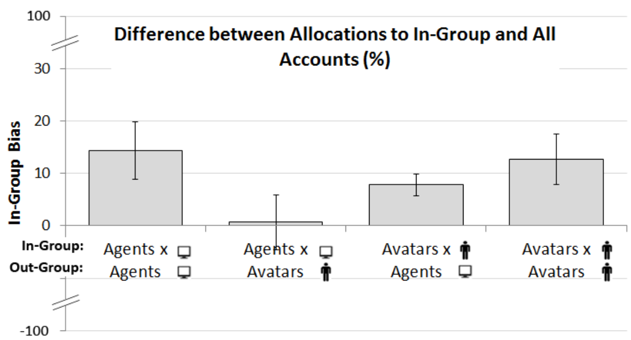

Figure 3: In-group bias in Experiment 3. The error bars show the standard errors.

Experiment 4

The previous experiments showed that it is possible to compensate for the fact that computers do not belong to the human social category by changing the computer's visual appearance (Experiment 2) or the structure of the task (Experiment 3). In Experiment 4 we wanted to test if it was possible to over-compensate and have people offer more to computers than to humans. To accomplish this we had participants engage, in a between-participants design, in the nested social dilemma with an in-group that was always composed of agents of the same ethnicity as the participant but, with an out-group that was composed of avatars of either the same or a different ethnicity than that of the participant. For the case where both the in-group agents and out-group avatars had the same ethnicity, we expected to replicate the result in the previous experiment, i.e., no preference between the in- and out-groups. For the case where the out-group was composed of avatars of a different ethnicity, we expected people to favor the in-group agents. The rationale is that in this case two categories (ethnicity and payoff-defined group membership) favored agents and only one favored avatars (human category). We recruited 47 participants on Amazon Mechanical Turk. Most participants were Caucasian (25.5%) or East Indian (59.6%). The procedure to support the deception regarding avatars was similar to the one used in Experiment 1.

To facilitate interpretation, we converted ticket allocations into percentages. As in the previous experiment, we use the difference between allocations to the in-group and all accounts as a measure of the in-group bias. Means and confidence intervals for this measure are shown in Figure 4. We ran one-way *t*-tests to compare the differences in each condition to zero: if the difference is statistically significantly different from zero, then there is evidence for an in-group bias. The results confirmed our prediction: for in-group agents and out-group avatars of the same ethnicity, the in-group bias was not statistically significantly different than zero, $t(22) = .417$, $p = .681$, $r = .089$; however, for in-group agents of the same ethnicity and out-group avatars of a different ethnicity, the in-group bias was statistically significant, $t(23) = 2.13$, $p = .044$, $r = .406$.

The results, therefore, showed that by associating more positive social categories with computers than with humans, it is possible to overcome people's bias in favor of humans.

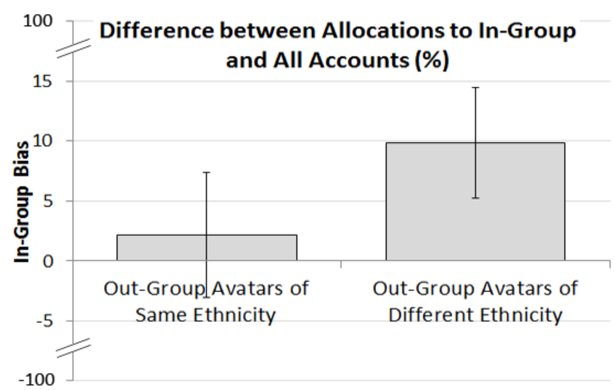

Figure 4: In-group bias in Experiment 4. The error bars show the standard errors.

Discussion

As computational systems take an active role in today's society, it becomes important to understand how people reach decisions with non-human decision makers. Social categorization theory (Tajfel & Turner, 1986; Turner et al., 1994; Crisp & Hewstone, 2007) provides important insights. Accordingly, we argue that the same mechanism whereby people form social categories that include some people (the in-groups) and exclude others (the out-groups) extends to computers. At first blush, people categorize humans as being in-group members, in the human social category, and computers as the out-group. This differentiation is motivated by perceptions that humans possess more sophisticated mental abilities than computers (Blascovich et al., 2002; Gray et al., 2007; Waytz et al., 2010). Consequently, people tend to favor humans over computers in resource allocation (Haslam, 2006; Loughnan & Haslam, 2007). However, by manipulating the characteristics of artificial decision makers (e.g., their ethnicity) or of the situation (e.g., payoff interdependence) it is possible to bring other social categories into play, which can compensate (and even over-compensate) for people's bias in favor of humans.

To support our proposal we presented four experiments where participants engaged with computers perceived to be controlled by algorithms (agents) or by humans (avatars). Experiment 1 showed that people offered in the dictator game more money to avatars than to agents, supporting the existence of a bias in favor of humans. Experiment 2 introduced a second social category – the counterpart's ethnicity – and data indicated that people offered more money to any counterpart, avatar or agent, that had the same ethnicity. Experiment 2 also indicated that the effects of the human and ethnicity social categories can be independent and additive, thus suggesting that the human/computer categorization is another in the long list of categories that people use in social life (Crisp & Hewstone, 2007). Experiment 3, in turn, demonstrated that it is also possible to promote group membership with computers by creating structural interdependence, based on shared payoffs. The results showed that in the nested social dilemma when the in-group was composed of agents and the out-group of avatars, people would not favor avatars anymore. Finally, Experiment 4 showed further that people can favor agents to avatars if the former are associated with more positive categories than the latter. Effectively, people offered more in the nested social dilemma to in-group agents of the same ethnicity than to out-group avatars of a different ethnicity.

These results have implications for understanding people's behavior with computers. The "computers are social actors" theory (Reeves & Nass, 1996) introduced the idea that people can treat computers in a fundamentally social manner. Indeed our results demonstrated that, even in the absence of financial incentives, people were willing to offer money to computers (Experiments 1 and 2). However, our results complement this view by demonstrating that there are still important differences in the way people treat computers in social settings, when compared to humans. Everything else being equal, people tended to favor humans (the in-group) to computers (the out-group). The results suggest that social categorization is a useful framework to understand people's decision-making behavior with computers.

The results comport with findings that people naturally attribute more mind to humans than computers (Blascovich et al., 2002; Gray et al., 2007; Waytz et al., 2010). The expectation of less mental abilities could be the fundamental reason people fail to treat computers as in-group members and, consequently, show a bias in favor of humans. Future work should, therefore, test the prediction that proper simulation of appropriate mental abilities suffices to make people treat computers in the same manner as humans, at least in the context of decision-making tasks with clear financial incentives.

This work also presents further evidence that people have the cognitive resources and motivation to classify themselves and others along multiple categories simultaneously (Crisp & Hewstone, 2007). Our results show clear evidence for an additive pattern that reflects a positive correlation between group differentiation and intergroup bias. However, researchers have also pointed out that, for people that strongly identify with the group, this correlation can be negative, i.e., the less the differentiation, the higher the bias (Brewer, 1991). The rationale is that bringing groups together via shared categories might threaten people's desire for distinctiveness. Future work, thus, should explore if in-group identification can also moderate people's bias in favor of humans when engaging with computers in other settings.

From a practical point of view, this work emphasizes the importance of considering appropriate social and cognitive psychological theories of human behavior when designing artificially intelligent decision makers. It is important designers understand and compensate for people's tendency to reach different decisions according to whether they perceive computers to be driven by a human or by computer algorithms. Superficially, designers could try to de-emphasize that certain decisions are being made autonomously; however, there are ethical and legal concerns that might 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. Looking instead to the vast literature on intergroup conflict resolution suggests a set of more-principled design guidelines. First, personalizing out-group members and increasing intergroup contact can reduce in-group bias and prejudice (Pettigrew & Tropp, 2006). In the case of artificial decision makers, this suggests that designers should strive to increase visibility and transparency for the mechanisms and reasons behind the decisions that were made. Second, creating cooperation through shared goals is known to reduce intergroup competition (Sherif et al., 1961). This suggests, for instance, emphasizing that the policies that decision algorithms implement serve the common good or creating payoff interdependence between humans and computers. Finally, in case of severe conflict, designers can resort to negotiation techniques (Pruitt & Carnevale, 1993) to resolve divergence of interest between humans and computers, an approach that is in fact an active topic of research (Lin & Kraus, 2010).

A multi-disciplinary perspective is critical for understanding how people adapt to changing conditions in an evolving world and, in particular, carry the mechanisms of human-human interaction into human-computer interaction. With distinct cooperation tasks and with different kinds of populations, we showed that the mechanisms of social categorization and intergroup behavior can explain people's interaction with computers that make decisions. Future work should further explore more decision contexts, more social categories (e.g., age, gender, culture), more roles (e.g., receiver), other kinds of machines (e.g., robots), and determine the sufficient conditions artificial decision makers should possess in order to be treated in the same manner as human decision makers.

Acknowledgments

This research was supported in part by the following grants: FA9550-09-1-0507 from the Air Force Office of Scientific Research, IIS-1211064 and SES-0836004 from the National Science Foundation. The content does not necessarily reflect the position or the policy of any Government, and no official endorsement should be inferred.

References

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*, 103-124.

Brewer, M. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475-482.

Crisp, R., & Hewstone, M. (2007). Multiple social categorization. *Advances in Experimental Social Psychology*, *39*,163-254.

Davenport, T., & Harris, J. (2005) Automated decision making comes of age. *Sloan Management Review*, *46*, 83-89.

Dawes, R (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169-193.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*, 583-610.

Ford, R. (2008). Is racial prejudice declining in Britain? *British Journal of Sociology*, *59*, 609-634.

Forsythe, R., Horowitz, J., Savin, N., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*, 347-369.

Gaertner, S., & Dovidio, J. (2005). Understanding and addressing contemporary racism: Aversive racism to the common intergroup identity model. *Journal of Social Issues*, *61*, 615-639.

Gallagher, H., Anthony, J., Roepstorff, A., & Frith, C. 2002. Imaging the intentional stance in a competitive game. *NeuroImage*, *16*, 814-821.

Gray, H., Gray, K, & Wegner, D. 2007. Dimensions of mind perception. *Science*, *315*, 619.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252-264.

Honignam, B. (2012). 100 Fascinating social media statistics and figures from 2012. *The Huffington Post*. Retrieved from: http://www.huffingtonpost.com/brian-honigman/100-fascinating-social-me_b_2185281.html.

Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE*, *3*, 1-11.

Lin, R., & Kraus, S. (2010). Can automated agents proficiently negotiate with humans? *Communications of the ACM*, *53*, 78-88.

Loughnan, S., & Haslam, N. (2007). Animals and androids: Implicit associations between social categories and nonhumans. *Psychological Science*, *18*, 116-121.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. National Academy of Sciences*, *98*, 11832-11835.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*, 81-103.

Nass, C., Fogg, J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*, 669-678.

Pettigrew, T., & Tropp, L. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality Social Psychology*, *90*, 751-783.

Pruitt, D., & Carnevale, P. (1993). *Negotiation in social conflict*. (Brooks/Cole, Pacific Grove, CA).

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places* (Cambridge University Press, New York).

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G. & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, *35*, 395-405.

Rossen, B., Johnsen, K., Deladisma, A., Lind, S., & Lok, B. (2008). Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases. In *Proceedings of the Intelligent Virtual Agents Conference*, pp. 237-244.

Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., & Cohen, J. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755-1758.

Sherif, M., Harvey, O., White, B., Hood, W., & Sherif, C. (1961). *Intergroup conflict and cooperation: The Robbers Cave experiment*. Norman, OK: University Book Exchange.

Shih, M., Pittinsky, T., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80-83.

Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp 7-24). Chicago: Nelson-Hall.

Turner, J., Oakes, P., Haslam, S., & McGarty, C. (1994) Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454-465.

U.S. Census Bureau. (2011). E-stats - Measuring the electronic economy. Retrieved from: http://www.census.gov/econ/estats/2011reportfinal.pdf.

Waytz, A., Gray, K., Epley, N., & Wegner, D. (2010). Causes and consequences of mind perception. *Trends in Cognitive Science*, *14*, 383-388.

Wit, A., & Kerr, N. (2002). “Me versus just us versus us all” categorization and cooperation in nested social dilemmas. *Journal of Personality and Social Psychology*, *3*, 616-637.