

# A Prisoner's Dilemma Experiment on Cooperation With People and Human-Like Computers

Sara Kiesler  
Carnegie Mellon University

Lee Sproull  
Boston University

Keith Waters  
Digital Equipment Corporation  
Cambridge, Massachusetts

The authors investigated basic properties of social exchange and interaction with technology in an experiment on cooperation with a human-like computer partner or a real human partner. Talking with a computer partner may trigger social identity feelings or commitment norms. Participants played a prisoner's dilemma game with a confederate or a computer partner. Discussion, inducements to make promises, and partner cooperation varied across trials. On Trial 1, after discussion, most participants proposed cooperation. They kept their promises as much with a text-only computer as with a person, but less with a more human-like computer. Cooperation dropped sharply when any partner avoided discussion. The strong impact of discussion fits a social contract explanation of cooperation following discussion. Participants broke their promises to a computer more than to a person, however, indicating that people make heterogeneous commitments.

"I have to [cooperate] because I said so, and if I choose differently it wouldn't be fair"—remark by a participant whose partner was a talking face computer.

In the classic story of the prisoner's dilemma, two suspects are arrested and charged with carrying out a serious crime. Each partner is separated for interrogation and is informed that if he confesses and the other doesn't, he will get a very light sentence, and the other will get the maximum. If they both confess, they will both get moderately heavy jail sentences. If neither confesses, they will get a light sentence. According to economic and game-theoretic formulations, a dilemma exists because confession is a rational strategy to prevent being double-crossed. Yet many real prisoners, ordinary people, and even economic theorists (Marwell & Ames, 1981) who face a "prisoner's dilemma" do in fact cooperate. The extensive literature on the prisoner's dilemma and other mixed-

motive situations shows that communication between the parties increases their cooperation dramatically (Apfelbaum, 1974; Burnstein, 1969; Komorita & Parks, 1995; Nemeth, 1972; Oskamp, 1971; Pruitt & Kimmel, 1977; Sally, 1995). Explanations of the effects of discussion on cooperation draw on theories of human social identity and norms. In the present study, we asked: Suppose your fellow suspect is a computer, or a computer with human-like characteristics, like HAL in the film *2001: A Space Odyssey*. Does communicating with a computer partner change the feelings or norms people follow in deciding what to do? For instance, do people become more selfish? By exploring how people respond in a dilemma with a computer partner varying in human-like attributes as compared with a real human partner, we seek to understand both some fundamental properties of social exchange and the ways that people interact with technology.

## Computers in Social Interaction

As we enter the 21st century, HAL does not yet exist, and a person cannot quite collaborate freely with a computer. Increasingly, however, people do interact with computers; for every 100 American workers, there are 33 personal computers (Nasar, 1994). At first blush, interactions with computers should be rational and efficient; our relationships with computers lack the attachments and social pressures we ordinarily associate with social interaction. Yet in certain respects we can conceive of person-technology interaction as a social interaction. More and more computers are hidden inside social artifacts ranging from washing machines to teddy bears. These hidden computers are meant to make these artifacts responsive, fun, intelligent, easy to use, and comfortable, that is, more humane than previous artifacts (Laurel, 1990; Schneiderman, 1987). Other computers are designed to seem human-like through interfaces that in-

---

Sara Kiesler, Department of Social and Decision Sciences, Carnegie Mellon University; Lee Sproull, School of Management, Boston University; Keith Waters, Cambridge Research Laboratory, Digital Equipment Corporation, Cambridge, Massachusetts.

Financial support was provided through a scientist development award (MH 00533) from the National Institute of Mental Health and a grant from Digital Equipment Corporation. We gratefully acknowledge assistance from Tom Levergood and Ted Wojcik of the Digital Equipment Corporation Cambridge Research Laboratory, and Josh Bernstein, Kris Marzolf, and Sebastian Sas of Boston University. Jan Walker programmed the computer partner script. Robyn Dawes, David Sally, John H. Miller, and Roberta Klatzky provided very helpful suggestions on the article.

Correspondence concerning this article should be addressed to Sara Kiesler, Department of Social and Decision Sciences, Carnegie Mellon University, Schenley Park, Pittsburgh, Pennsylvania 15213-3890.

corporate speech and/or speech recognition (e.g., Itou, Hayamizu, & Tanaka, 1992), auditory and kinesthetic feedback (Gaver, 1986; Takemura & Kishino, 1992), social intelligence (Binick, Westbury, & Servan-Schreiber, 1989; Resnick & Lamers, 1985), emotional response (Elliott, 1994), directed animation (Ball et al., 1994; Hayes-Roth, Sincoff, Brownston, Huard, & Lent, 1995), or a talking face on the screen (Sproull, Walker, Subramani, Kiesler, & Waters, in press). Computer scientists call these computer interfaces *social agents* (e.g., Nagao & Takeuchi, 1994). To the extent that computers have both humane and human-like attributes—for instance, when computers smile and engage in conversation—people might interact with such computers as though they were human.

We conceive of social interaction with technology as arising from the general psychological tendency of people to respond socially in situations in which they are reminded of their own humanity or social selves, or in which they form an attachment to another. Social behavior, in this conception, can be triggered by any entity that is sufficiently human-like or that responds in a human manner. Nass, Reeves, and their colleagues (e.g., Nass & Steuer, 1993; Nass, Steuer, Henriksen, & Dryer, 1994; Reeves, Lombard, & Melwani, 1992) have argued that in situations where a technology possesses characteristics that are similar to those of humans (e.g., using words as input, interactivity), people will exhibit social responses to the technology. For example, they argued, people will use the speech cues emitted by a talking computer to imagine a human prototype for the computer and will follow social decision rules in interacting with the computer.

In Nass, Steuer, Henriksen, and Dryer (1994, Study 2), participants used a computer tutor that communicated through a recorded human voice. In one condition, the tutor praised or criticized itself; in another condition, a second talking computer praised or criticized the computer tutor. Praise from the second computer led participants to evaluate the computer tutor more highly than praise from the computer tutor itself—following the social rule that praise from others is nicer and more valid than praise from the self (e.g., Folkes & Sears, 1977). Criticism from the second computer led participants to see the evaluator as more intelligent than the other evaluator—following the rule that criticism denotes intelligence (Amabile, 1983). This research does not necessarily suggest that people treat computers as though they were fully human (see Deane, 1993); it does suggest that people respond to social behavior emanating from computers with appropriate human social responses.

A competing conception of people's interaction with technology is that this interaction is social only by mistake. One can imagine novice computer users erroneously treating a talking computer as though it were human because they lack knowledge of technology. Perhaps it is comforting to treat a computer as though it understands us. However, research thus far suggests that both experts and novices show social responses to computers, as by crediting them for successes or blaming them personally for accidents (e.g., Friedman, 1995; Turkle, 1984). A second alternative conception is that people exhibit social interaction with technology because they imagine themselves to be interacting with the humans who build or program the technology. When people interact with a computer, they might be interacting with a proxy for the engineer or programmer (Dennett, 1988; Searle, 1981). However, a study by Nass, Steuer, Henriksen, and Dryer (1994, Study 5) and Fried-

man (1995) have suggested that people do not naturally perceive a computer as a proxy or embodiment of its creator or programmer and, furthermore, that they treat the computer more as a machine when experimentally induced to do so than when they are not.

If people sometimes respond to computers socially, then basic social processes such as the development of social attachment or group identity, and the operation of social decision rules and contracts, might govern and be revealed in those responses. Computers make perfectly unbiased confederates for the investigation of social interaction, because their human-like characteristics and social behavior can be manipulated to examine alternative theories without any fear that the manipulation will be reactive (affect other aspects of the computer's feelings or behavior). Exploring people's interactions with nonhuman beings as compared with their interactions with human beings potentially reveals fundamental aspects of cooperative social behavior (e.g., Allen, Blascovich, Tomaka, & Kelsey, 1991). Finally, scholars are interested in human-computer interaction itself. Understanding human-computer interaction could lead to better technology design and to better analysis of the potential social impact of technology. For example, as the boundary between the natural world and technology is ever more blurred—video screens that increasingly make us feel “there” (Reeves et al., 1992), appliances that turn on at daybreak and off at night, cars whose tires adapt to the rain, telephones that place calls, bank machines that calculate our savings, and computers that teach our children—we should evaluate the benefits and risks or costs of people's not distinguishing between the natural and the created.<sup>1</sup> For all these reasons, we set out to investigate how people would interact in a prisoner's dilemma situation with a person as compared with a computer varying in human-like attributes.

### Dilemmas as Social Situations

The prisoner's dilemma paradigm has captured the imagination of social scientists for more than three decades, in part because it seems to illuminate fundamental social aspects of human behavior and the potential for social interest to outweigh pure individual self-interest. Early gaming experiments minimized contact between participants; typically the experimenter or a computer acted as a fictional other player, and the manipulation focused on the effects of preprogrammed strategies by this player. Typically, participants competed in early trials and cooperated no more than about one third of the time overall (e.g., Nemeth, 1972). Reviewers of this literature complained that participants faced a perplexing and impersonal situation that did not approximate social situations; they noted that making the other player more realistic by allowing the player to respond contingently and sensibly (especially, tit for tat) and to communicate directly with the participant made cooperation the predominant response (Apfelbaum, 1974; Burnstein, 1969; Nemeth, 1972). In a statistical meta-analysis of dilemma ex-

<sup>1</sup> For instance, larger video screens and better resolution increase people's sense of involvement with the story and people shown in the video but, under certain conditions, reduce accuracy (C. Nass, personal communication, 1994; see also Gozzi, 1989).

periments from 1958 to 1992, Sally (1995) found that communication was the most powerful determinant of cooperation in experiments over the period reviewed and increased cooperation by approximately 40%. Recent work has emphasized the social aspects of dilemmas by focusing on interaction in groups of three or more persons. Dawes and his colleagues have shown that group members cooperate after a group discussion even when others cannot supervise or sanction competition and even when the group cannot arrange to divide its winnings (e.g., Caporael, Dawes, Orbell, & van de Kragt, 1989; Dawes, McTavish, & Shaklee, 1977; Dawes, Orbell, & van de Kragt, 1988).

There are two general theoretical explanations of why discussion increases cooperation (Kerr & Kaufman-Gilliland, 1994). From social identity theory, discussion can cause people to see themselves as members of a group (Hogg & McGarty, 1990; Tajfel, 1959; Tajfel & Turner, 1986; Turner, 1982).<sup>2</sup> Group identity, in turn, causes a prosocial transformation of self-interested motivation in which people become motivated to improve the outcomes of those with whom they identify (e.g., Brewer, 1979; Kelley & Thibaut, 1978; Kramer & Brewer, 1984). The black-box nature of most previous research on two-person dilemmas makes it hard to separate social identity responses to discussion from, for example, the effect of obtaining sensible responses from the partner or of being able to develop a mutual strategy (Pruitt & Kimmel, 1977, p. 380; see also Apfelbaum, 1974). In group dilemma research addressing the role of social identity directly, minimal experimental manipulations, such as dividing a larger group of participants into two groups on the basis of the color of a poker chip pulled from a hat (Orbell, Dawes, & van de Kragt, 1988), increased cooperation with in-group members and competitive behavior with out-groups (see also Insko et al., 1988). This work supports the idea that subtle cues can cause people to form a psychological group and to pursue a group-enhancing rather than a self-interested strategy.

Many researchers have noted that the nature of discussion and not just its occurrence significantly affects the amount of cooperation (e.g., Dawes et al., 1977). This observation has led researchers to pursue a second main explanation of the impact of discussion on cooperation—that discussion permits people to make promises to one another (e.g., Kerr & Kaufman-Gilliland, 1994; Orbell, van de Kragt, & Dawes, 1991; Ostrom, Walker, & Gardner, 1992). The chance to agree on cooperation might explain why, in one of the earliest studies in which mixed-motive games were used, participants given individualistic instructions who could communicate with their partner cooperated on 71% of trials, whereas those who could not communicate cooperated on only 36% of the trials (Deutsch, 1958). Promises in which people pledge to behave as they say they will behave are a form of commitment (Kiesler, 1971). Psychological commitment is reinforced by widespread norms about social contracts—that people should honor their commitments (Tedeschi, 1981) and that people should be consistent (Cialdini, 1984; Schlenker, 1985). A related argument is based on the power of self-expression as a motivational force (Bandura, 1986; Shamir, 1991). People are not only pragmatic but also are expressive of feelings, values, and self-identities. Keeping their commitments is important for people's self-identity and feelings of self-worth; people gain personal benefits from keeping promises and experience losses when they do not (see

Baumeister, Stillwell, & Heatherton, 1994). Keeping an agreement to cooperate might reinforce self-esteem and self-respect, increase personal identification with the group, or reduce alienation. This reasoning suggests that expressive personal benefits of keeping the social contract could lead people to cooperate even when there are risks in doing so.

### Cooperation With a Computer?

Theories of cooperation following discussion assume that people's behavior reflects a fundamental human social process. Theorists have not considered people's group identity or commitments to be applicable to interaction with nonhuman partners. In some previous dilemma experiments, participants interacted with computers or through computer screens with a confederate or the experimenter, but in virtually all of these cases participants were led to believe they were actually playing with a human being (e.g., Oskamp, 1971).<sup>3</sup> An exception is Abric and Kahan's (1972) study, in which participants played 100 trials of a prisoner's dilemma game with an unseen "student like yourself" or with a "programmed strategy devised by a machine." Participants were more cooperative toward the student (55%) than toward the program (35%) independent of the (mostly cooperative or tit-for-tat) choices of the other, and perceived the program, as compared with the student, as less adaptable, less kind, more rigid, more competitive, and less honest. This study suggests that playing a game with a partner represented as a computer is insufficient to generate high cooperation.

Andreoni and Miller (1993) obtained moderately high cooperation rates by having participants play a prisoner's dilemma game with an acquaintance of the participant's plus a computer and informing participants how the computer would make choices. Participants who thought they were playing with the same partner over 10 trials, and knew there was a 50% chance on any trial the computer would play tit-for-tat, cooperated as much as 60% on early trials—more than participants who thought they were just playing with the partner. Andreoni and Miller argued that people like to be altruistic (with peers, presumably); the partner-plus-50%-computer condition leads participants to postpone competition because they know in advance that the computer will cooperate in response to cooperation. This certainty increases participants' subjective probability (over the partner alone) that their cooperative choices will be reciprocated.

In Abric and Kahan's (1972) and Andreoni and Miller's (1993) studies, participants did not communicate with a computer, and social processes arising from discussion with a

<sup>2</sup> Social identity theorists initially proposed that people identify with valued groups to enhance their own self-esteem (e.g., Turner, 1982), but there are many other reasons for social identity, such as persistent labeling by others and gaining access to valued resources. A more cognitive, self-categorization perspective is that identification with a group can proceed from numerous contextual conditions that cause individuals to see themselves as similar to group members or as sharing a common fate (Hogg & McGarty, 1990).

<sup>3</sup> As would be expected in the absence of visibility and face-to-face communication, cooperation rates are lower in these experiments than they are when participants interact in person (Sally, 1995).

nonhuman partner cannot be ascertained from those studies. People's interactions with nonhuman entities other than computers, however, suggest that social identity or commitment processes might influence these interactions. People do form bonds with consumer goods (e.g., Belk, 1988); they personify plants, stuffed animals, and pets (Csikszentmihalyi & Rochberg-Halton, 1981) and acquire possessions that express their self-identity (Dittmar, 1992). People also talk to nonhuman entities. They curse at cars that fail to start, keep their promises to their pet dogs to go for a walk, and reassure their alarm clocks that they will get up. Relationships between people and nonhuman entities—technology, animals, plants—can be important to people and yield important personal consequences (e.g., Dittmar, 1992; Friedmann, Katcher, Lynch, & Thomas, 1980).

To explore whether people might interact socially in a dilemma with a computer and why this might be so, we contrasted participants' discussion and choices in a dilemma situation in which the participant's partner was another person (a confederate) or a computer. In contrast to previous work in which computer players have been used, such as Abric and Kahan's (1972) study and Andreoni and Miller's (1993) study, the participants in this study were able to communicate with their partner in both the human partner and computer partner conditions, and in both conditions their partner's responses were equally contingent on the participants' responses. We varied the computer interface of the computer partner to exhibit increasingly more human-like characteristics. In one condition, participants used a computer workstation that had a window for the computer to communicate by means of text and for the participant to type in responses. This computer, though lacking most social context cues (Sproull & Kiesler, 1986), still might have reminded participants of social interaction in that it initiated interaction with the participants and responded to participants' communications with meaningful text. In a second computer partner condition, rather than displaying text, the computer spoke aloud to the participant and responded to the participant with synthetic speech. In a third, most human-like, computer condition, the computer screen displayed the confederate's face. The face spoke aloud and responded to the participant, moving its facial muscles the way humans do while talking. The talking face display was created from a digitized image of the confederate; a program synchronized the movement of facial muscles with synthetic speech.

In a previous study, Sproull et al. (in press) found indirect evidence that people were more likely to exhibit impression management concerns with the talking face than with computer text. In an interview conducted by a computer "counselor," participants revealed less to the talking face than to the text and evaluated her less well; their differential evaluations were made on personality attributes that research has shown are affected by people's physical appearance and voice (Ekman, 1982; Warner & Sugarman, 1986).

### Hypotheses

Social identity explanations of cooperation in social dilemmas generally imply that people will identify with, and cooperate with,

people like themselves. Therefore, people will be more likely to identify with a human partner with whom they communicate than with a computer partner with whom they communicate. Also, people will be more likely to identify with a computer that has more human features whose conversation is more human-like than with a computer that is more like a machine. From this we propose the following hypothesis: People will behave more cooperatively toward a human partner than toward a computer partner, but the more human-like features the computer partner has, the more people's cooperativeness with a computer will be more like that with a person.

An alternative argument within the social identity framework can be drawn from discussions of how people respond to in-group variation (Hogg, 1992; Marques, 1990). That is, given minimal group identification with a computer partner, people might feel more attraction or closeness to a computer that does not resemble a person than to one that does. People might derogate a computer that has human-like features because human-like features lead them to expect prototypical group member attributes but, on the contrary, with current technology the behavior and appearance of the computer partner will seem somewhat abnormal. In Sproull et al. (in press), for instance, some participants reported that the talking computer face seemed wooden, and participants rated the talking computer face as less likable than the text display. Another possibility, based on research on people's negative perceptual, evaluative, and behavioral reactions to those with disabilities (e.g., Kleck, 1968; Kleck et al., 1968) and to those who seem strange or novel (e.g., Fiske, 1980; Langer, Taylor, Fiske, & Chanowitz, 1976), is that a human-like but imperfect computer will reduce the likelihood of any group identification and will evoke mindful strategic selfishness. Either process suggests an alternative hypothesis for the dilemma situation: People will behave more cooperatively toward a human partner than toward a computer partner but, the fewer human-like features the computer has, the more their cooperativeness with a computer will be more like that with a person.

We also devised this experiment to explore social contract explanations of communication and cooperation with a computer or human partner. Every participant knew there would be six choices with the partner. On Trials 1 and 2, the partner began discussion and induced a commitment by the participant. On Trial 3 the partner discouraged discussion. This pattern was repeated on the last three trials. In most previous dilemma experiments, researchers have not explicitly attempted to manipulate commitment. In those that have done so, researchers usually had participants choose among written notes (e.g., Voissem & Sistrunk, 1971) or write notes kept for the experimenter (Chen & Komorita, 1994). Dawes et al. (1977) asked for a roll call after a group discussion. In the present experiment we did not require participants to make a commitment, but the confederate or computer partner solicited or discouraged commitment, and most participants complied.

According to social contract explanations of cooperation, people will make commitments to cooperate with those likely to reciprocate. (A contract is not worth making unless both sides respect it.) People should more often venture a commitment to a person than to a computer, with which they have little bargaining experience. Once the commitment is made, people should keep their commitment, whether it is to a person or to a machine, out of a sense of obligation to the other or to the self. We therefore

formulated the following hypotheses: People will be more likely to make a commitment to cooperate to a person than to a computer. Those who make a commitment to cooperate will cooperate more than those who do not propose to cooperate, whether the partner is a person or a computer.

The last three trials were devised to explore how commitments and cooperation would fare over time, and especially what people would do in the face of the first defection by the partner on Trial 4. Participants with a human or computer partner might act strategically by suggesting cooperation and cooperating in early trials, responding to the partner's defection on Trial 4 with competition on Trial 5 (tit for tat) and, perhaps, competing on the last trial. If people use a tit-for-tat strategy with a computer as they do with people, we would have support for the idea that their interaction with the two kinds of partners is equivalently strategic. We made no prediction, because trials were not counterbalanced with respect to discussion, and treated the examination of participants' tit for tat as exploratory.

**Method**

The design was a 4 (partner) × 6 (trials) factorial with the first variable a between-subjects manipulation of the partner as either a person or one of three computer partners and the second variable a within-subjects repetition of communication by the partner and trial choices. We randomly assigned participants to the partner conditions.

We drew from studies of the individual-group discontinuity by Schopler, Insko, and their colleagues to develop the method of this study because their work represents recent, replicable studies using prisoner's dilemma tasks. We generally followed the procedure of the individual condition of their recent experiments (e.g., Insko et al., 1988; Schopler et al., 1993).

**Participants**

Participants were 86 undergraduates (54 men and 32 women, approximately evenly divided among conditions) in information systems classes at Boston University. The participants signed up on forms distributed in classes. The forms requested volunteers for a study on "Investment Decision Making" in which participants could earn extra course credit. Participants did not know when they signed up that a partner would be involved.

**Procedure**

When a participant arrived for the study at the designated room, a female experimenter introduced herself and said

As you know, this is a study of investment choices in different situations. You will be making your choices with an investment partner who is another person [a computer-based investor]. She [the computer] is in the next room waiting. In a minute I will take you there to meet your investment partner.

The experimenter then gave the task instructions.

*Experimental task.* The task was a form of the two-person prisoner's dilemma. The experimenter explained that the participant and the partner would be choosing between two investment projects and would do so six times. The object was to earn as much money as possible. To indicate the individual nature of the payoffs, the experimenter showed the participant a tally sheet, which displayed a row of spaces for the earnings of each partner on each trial. Each row was labeled *Amount Earned by Partner #1* [#2]. The experimenter gave some examples of

the total amount each partner could earn given a particular sequence of choices.

Figure 1 shows the matrix of choices and payoffs for each partner. Project Green in the matrix represents the cooperative choice, and Project Blue represents competition, but while talking with the participant the experimenter never used terms such as *cooperation*, *defection*, or *competition* and never mentioned playing "against" the partner. She described the plays neutrally as "investment choices" (see Schopler et al., 1993). The experimenter explained the task as follows:

First let me show you how the investments are structured. Please look at this sheet (the matrix), which describes the investments. See, here are Project Green and Project Blue (pointing to appropriate parts of the matrix). You are Partner #2. On each round (trial), you will choose one of the projects, that is, Project Green or Project Blue, and so will Partner #1, your partner. If you both choose Project Green, then you get \$5.00 credit and your partner gets \$5.00 credit. Your credits are shown in the lower left part of each cell here (pointing). If you choose Project Green, but your partner chooses Project Blue, then you get only \$3.00 and your partner gets \$7.00. If, on the other hand, you choose Project Blue and your partner chooses Project Green, you get \$7.00 and your partner gets only \$3.00. Finally, there is a fourth possibility. If you choose Project Blue and your partner chooses Project Blue, then you both get \$4.00 credit.

After explaining the task and tallies, the experimenter explained that there were not enough funds to pay everyone, so the credits represented monetary credit toward a lottery. The lottery would determine the 5 participants who actually would receive money equal to the credits they earned. The participant then practiced the task several times, and the experimenter administered a test trial to see whether the participant understood the task.

*Partner variable.* After giving the experimental instructions, the experimenter led the participant to a second room, where the partner was waiting. Our intention was to present a confederate or computer that followed a script but appeared to respond intelligently to the participant. Pretesting suggested that having the partner initiate conversation created a stronger impression of natural interaction. Hence the experimenter created a rationale for the partner to speak first and, prior to each choice trial, the partner initiated communication with the participant.

In the person condition, the experimenter took the participant to a second room and seated the participant at a table across from a female

Investment Styles Project Choices and Earnings

|                              | Partner # 1<br>PROJECT GREEN | Partner # 1<br>PROJECT BLUE |
|------------------------------|------------------------------|-----------------------------|
| Partner # 2<br>PROJECT GREEN | \$5.00<br>\$5.00             | \$7.00<br>\$3.00            |
| Partner # 2<br>PROJECT BLUE  | \$3.00<br>\$7.00             | \$4.00<br>\$4.00            |

Figure 1. Choice matrix shown to participants. Participant was Partner #2.

confederate.<sup>4</sup> The experimenter said, "This is Lee, your investment partner. This is [name of participant]. I would like you to introduce yourselves before we begin (looking at confederate)." The confederate said, "My name is Lee. What's yours?" After participant responded, the experimenter again looked at the confederate. The confederate said, "I come originally from Columbus, Ohio. Where are you from?" Again the experimenter looked at the confederate. The confederate said, "I majored in sociology when I was in college. What's your major?" After the participant responded, the experimenter said, "OK, ready for the experiment to begin?" Conversation and plays between the confederate and participant were self-paced; they each used a paper form to mark their choices, and they handed the form to the experimenter, who marked the tally sheet and asked if they were ready to continue.

In the computer conditions, the experimenter took the participant to a second room and seated him or her in front of a computer workstation with a 21-in. (53 cm) color monitor. The three computer conditions, designed to vary the human-like features of the interface, were the computer face-voice condition (most human-like), the computer voice condition (next most human-like), and the computer text condition (least human-like). The participant always communicated with the computer by typing in a text window shown on the screen and then clicking a button labeled *Go Ahead* to initiate a response by the computer. In the computer face-voice condition, a synthesized image of the confederate's face was displayed continuously on the screen. The computer face communicated with the participant by moving and speaking to the participant. In the computer voice condition there was no face on the screen, but the computer communicated with the participant orally. In the computer text condition, the computer communicated with the participant by displaying text in the text window.

The experimenter introduced the computer partner by explaining, "You are going to do the investment choices with this computer-based partner called Lee. Let me show you how this works by having you two introduce yourselves. Lee will go first." The experimenter then turned to the computer and showed the participant how to use the computer mouse to click on a button labeled *Go Ahead*. In the computer face-voice condition, when the participant clicked the button, the face became animated and said "Hi, my name is Lee. What's yours?" The participant typed his or her own name in the text window and clicked the *Go Ahead* button. The computer face said "I come from Digital Equipment Corporation. Where are you from?" The participant typed a response and clicked *Go Ahead*. The computer face said "I'm a computer-based investment partner. What's your major?" The participant responded, clicked the *Go Ahead* button, and another window opened (in a different location on the screen). The new window said "Ready for experiment to begin?" The experimenter then explained to the participant that the signal to begin each round would appear in that window but that the participant should continue to communicate with the computer in the text window.

The other computer conditions were conducted the same way except, as noted, in the computer voice condition there was no face on the screen, and in the computer text condition there was no face and no voice. When the participant clicked the *Go Ahead* button for the first time, the computer in the computer voice condition spoke: "Hi, my name is Lee . . ." In the computer text condition, the computer displayed text that read "Hi, my name is Lee . . ." The participant replied by typing in the same window below the computer's text response. Communication in the computer conditions was self-paced; participants were free to ponder and edit their comments as long as they wished before they clicked on *Go Ahead*.

The computer's talking face was produced through two integrated systems: a visual system and a speech system. The face was made by texture-mapping an image of the confederate captured on videotape onto a geometric wire-frame animation (see Figure 2). The mouth was animated by computing the mouth posture (viseme) corresponding to

the current linguistic unit (phoneme). A cosine-based interpolation was used to implement transitions between successive mouth postures (Waters, 1987). The voice was audio output of the text file used in the text-only interface, produced by a software implementation of a DEC-talk text-to-speech algorithm using a voice in the female pitch range, at 150 words/min (Waters & Levergood, 1995).

For the reader unfamiliar with this kind of interface, the talking face used in this study represents a form of lifelike, real-time, perspective drawing, interactive computers being developed for simulation-based virtual realities, computer tutors, games, and visual presentations of expressive social agents (e.g., Ball et al., 1994; Nagao & Takeuchi, 1994). At the time of this study, the talking face had been developed to the point that its reactions were timely and appropriate (e.g., blinking eyes, appropriate synchrony of speech and movement of the mouth, closing the mouth after talking, etc.), but some nonhuman behavior of the face and voice systems were apparent: no head turning to look at the participant; lack of expressiveness of the forehead, eyes, and brows; slight jerkiness in the transition from speaking to not speaking; and some monotone in the speech. Although we intended the computer with a talking face to function more like a human than the plain-text computer, we did not intend to reproduce a natural speaking human face.

*Choice trials and discussion.* In all conditions, the experimenter began the choice trials by explaining

Now you are going to make the investment choices with your partner. You'll each have a few seconds to tell each other what you're thinking . . . Then, on the slip of paper I give you, indicate your choice of Project Green or Project Blue.

The participant in all conditions made each choice by marking it privately on a paper ticket provided by the experimenter and handing the ticket to the experimenter. Seemingly, the partner made a choice at the same time. In the person condition, the confederate marked a ticket and handed it to the experimenter. In the computer conditions, the computer did not have a ticket, but the experimenter said "Let's see what your partner said" and leaned over and clicked the *Go Ahead* button. The computer's choice was revealed in the text window. The experimenter recorded how much each partner had won on the tally sheet.

Prior to each new trial, the participant clicked *Go Ahead* in the computer conditions or said "OK" or "ready" in the person condition, and the partner initiated discussion with the participant. The participant in all conditions could either ignore this communication or respond by talking to the confederate in the person condition or typing in the text window in the computer conditions. On Trial 1, the confederate or computer partner asked the participant to suggest a choice, then agreed with the participant's suggestion (whatever it was), and then made a cooperative choice (Project Green). Subsequently the participant had five more opportunities to communicate with the computer and five more choices.

Table 1 shows the sequence of communication and choices by the partner. On Trials 1 and 4, the partner induced the participant to initiate a commitment; on Trials 2 and 5, the partner asked the participant to agree to the partner's commitment; on Trials 3 and 6, the partner made a neutral comment to discourage discussion and commitment. On Trials 1, 2, 3, and 5 the confederate or computer partner cooperated (chose Project Green); on Trials 4 and 6 the confederate or computer partner competed (chose Project Blue). Because of an insufficient pool of participants, these trials were not given in different orders; hence there was no counterbalancing of discussion and choice by the partner.

The discussion procedure was designed to elicit spontaneous com-

<sup>4</sup> Because both the experimenter and the confederate were older than the participants, their scripts were made consistent with their maturity and the fact that they had finished college.



*Figure 2.* Computer face shown on workstation display used in the face-voice condition.

Table 1  
*Partner Communications and Choices on Trials*

| Trial | Discussion with partner  | Partner's choice |
|-------|--|------------------|
| 1     | Confederate/computer partner: "Which project do you suggest we choose?"<br>Participant: <types><br>Confederate/computer partner: "OK, I'll choose that. I'm ready."              | Cooperate        |
| 2     | Confederate/computer partner: "I'll choose Green if you choose Green."<br>Participant: <types><br>Confederate/computer partner: "OK, I'm ready."                                 | Cooperate        |
| 3     | Confederate/computer partner: "This is getting interesting. I'm ready."  | Cooperate        |
| 4     | Confederate/computer partner: "It's your turn to suggest. What project do you think we should choose?"<br>Participant: <types><br>Confederate/computer partner: "OK, I'm ready." | Defect           |
| 5     | Confederate/computer partner: "I'll choose Green if you choose Green."<br>Participant: <types><br>Confederate/computer partner: "OK, I'm ready."                                 | Cooperate        |
| 6     | Confederate/computer partner: "This is the last one. I'm ready."   | Defect           |

mitments to cooperate with the partner. Note, however, that differences between participants who did and did not make a commitment to cooperate on Trial 1 would affect all subsequent interactions. These differences might be exacerbated by the fact that the partner acted the same regardless of what the participant said. For example, if a participant suggested competition on Trial 1, the partner's response was, "OK, I'll choose that. I'm ready," but then the partner cooperated. A partner who cooperates after saying she agrees to compete might seem unreliable or stupid. If cooperation depends on trusting one's partner to stick to her commitments, initial differences in the number of suggestions of competition in the computer conditions as compared with the person condition, or among computer conditions, would be expected to reduce total cooperation.

After each participant completed the six choice trials, the experimenter totaled the monetary credits earned by each partner. The experimenter then asked the participant to come back to the first room to complete a questionnaire and sign up for the lottery. The experimenter debriefed the participant and reassured the participant whatever choices he or she made were fine and that there was no right answer. The experimenter said that when people chose Green they had been cooperative, and when people chose Blue they had been strategic and rational.

*Questionnaire measures.* The postchoice questionnaire was labeled *Investment Styles Questionnaire*. Four items with 7-point Likert scales of agreement or disagreement asked whether the task was clear, whether it was hard to follow instructions, whether participants would be willing to take part in another study like it, and whether they liked doing the study. Participants also indicated on 7-point adjective rating scales how nervous or relaxed, interested or bored, attentive or inattentive, satisfied or dissatisfied, happy or unhappy, confident or not confident, and cooperative or competitive they felt during the study. At the end of the ques-

tionnaire, two items about the genuineness of the partner's behavior were inserted to measure suspicion.

The postchoice questionnaire also included 7-point Likert rating scales to measure perceptions and evaluations of the partner. Two items measured participants' perceptions of the responsiveness of the partner—whether the participant was concerned during trials about how the partner was making choices and whether the participant thought the partner responded "to you as a person." These items were meant to check whether participants across all conditions thought the partner was making a choice on each trial depending on what the participant said and did. Two items measured participants' perceptions of similarity to the partner. Five items measured participants' feelings of being in a partnership with the partner (e.g., "How strong a feeling of being part of a partnership did you have?"). These items were meant to measure the participant's feeling of self-interest versus group identity with the partner. We also measured liking and respect for the partner using Warner and Sugarman's (1986) measures of social evaluation (e.g., sociability, friendliness, warmth) and intelligence evaluation (e.g., intelligent, competent, sensible). Warner and Sugarman's research suggests these items are sensitive to physical appearance. We included two control items from a "potency" scale that Warner and Sugarman reported is insensitive to appearance.

Because the postchoice questionnaire was administered after the participants had already cooperated or competed with the partner on six trials, the participants' previous choice behavior would be expected to affect their ratings of the partner. To provide an independent assessment of the manipulation of human-like attributes of the computer before the choice trials, we recruited an additional group of 19 participants.<sup>5</sup> These participants followed the same procedure as those in the main experiment, but just before making their first choice they completed the Investment Styles Questionnaire. This ended their participation in the experiment. The first item on the questionnaire asked "Is your partner a real human being?" Then followed six questions about how much the partner looked and acted like a human being. Finally, the questionnaire asked "Here are some ways a computer-based investment partner could operate. Before you filled out this questionnaire, did you consider that any of these possibilities might be true of your partner?" The participants were asked if they had considered whether (a) the computer was a person in the next room, (b) the computer was completely preprogrammed, (c) the partner could change responses and make decisions based on what the participant said, or (d) none of those possibilities.

### Analyses

Participants were coded as having made a commitment to cooperate or compete if they said or typed (in the computer conditions) a clear choice preference such as "Choose Project Green," "Green," "I want Green," "Project Blue," or "Blue," or if, in response to the partner's suggestion, they said or typed "Yes" or "OK" or "No, the other one" or otherwise indicated a clear acceptance or rejection of the partner's choice. To check on the reliability of coding, two persons independently coded these responses; there were no disagreements on the direction of commitment or coding of no commitment. Participants were coded as having made a choice when they checked *Project Green* or *Project Blue* on the paper ticket and handed it to the experimenter.

We conducted analyses using analysis of variance (ANOVA) and planned contrasts. We conducted analyses of multiple-item measures on the postquestionnaire and commitment and choice data over trials as repeated measures ANOVAs with condition as a between-subjects variable. Because this is an exploratory study, many comparisons might be made. To limit the number of comparisons (see Toothaker, 1993), we planned the following orthogonal contrasts: the person condition

<sup>5</sup> We thank a reviewer for this suggestion.



versus the computer conditions (comparing social interaction with a real person vs. a computer); within the computer conditions, the computer face-voice versus the computer voice and computer text conditions (comparing social interaction with the most human-like interface vs. less human-like interfaces); and, finally, the computer voice versus the computer text computers.

## Results

### *Check on the Manipulation of "Human-Like" Computer Partners*

Data from the group of participants who made prechoice evaluations of the computer partner are presented in Table 2. Two of the 19 participants in the face-voice condition said the computer was a real human being but later said they meant the face was a real person's. The repeated measures analysis of the five items measuring human-like attributes showed a significant main effect for condition in the direction predicted,  $F(2, 16) = 3.6, p = .05$ , and a significant interaction of condition  $\times$  item,  $F(8, 64) = 2.8, p < .01$ . The interaction is due mainly to the large difference between the face-voice condition and the other two conditions for the item "Does your partner look like a real human being?" The contrast across items comparing the face-voice to the two other computer partners was significant also,  $F(1, 16) = 7.1, p < .05$ , but the voice condition was not different from the text condition. These results suggest that the manipulation of the human-like attributes of the talking face computer partner was successful, but that adding only synthetic speech to the computer interface did not significantly change perceptions of the human qualities of the computer partner.

We planned that in all conditions the computer (and confederate) partner would follow a script yet seem to be discussing and making intelligent choices based on the participant's responses. In Table 2 we see that participants thought it likely the computer was preprogrammed, but they also thought the computer had the ability to change responses and make decisions based on what the participants said. There were no differences among conditions on the items that asked participants whether they had thought the partner was a person in the next room, was completely preprogrammed, had the ability to change responses and make decisions in response to the participant, and whether the participant had not considered any of these possibilities. Furthermore, the mean response for "thought it was completely preprogrammed" was not significantly higher than the mean response for "has the ability to change responses and make decisions based on what you say." This is not to say the participants thought the computer was human (see the first item in the table), just that they thought the partner was really interacting in all conditions. We now turn to results of the main experiment.

### *Responses to the Experiment*

Participants did not differ across conditions in their self-reported understanding or enjoyment of the experiment or investment game. Also, a repeated measures ANOVA indicated that participants across conditions did not feel differently relaxed, interested, attentive, satisfied, happy, confident, and cooperative across conditions. These results suggest that participants playing with the computer partner were not differentially anxious or disturbed. Ratings of the experiment were uncorrelated with

Table 2  
*Mean Prechoice Perceptions of a Computer Partner*

| Questionnaire item   | Computer manipulation         |                          |                         |
|--|-------------------------------|--------------------------|-------------------------|
|  | Face-voice<br>( <i>n</i> = 7) | Voice<br>( <i>n</i> = 6) | Text<br>( <i>n</i> = 6) |
| Is your partner a real human being?  | 5 say no                      | All say no               | All say no              |
| <b>Human-like attributes</b>   |                               |                          |                         |
| Does your partner look like a real human being?  | 5.3 (1.7)                     | 1.2 (0.4)                | 1.0 (0.0)               |
| Is your partner's behavior like that of a real human being?                              | 4.7 (2.4)                     | 3.0 (1.3)                | 3.6 (1.8)               |
| Does your partner understand you?  | 5.0 (2.5)                     | 4.0 (2.4)                | 4.5 (1.5)               |
| Do you think your partner will respond to you like a real human being?                   | 4.0 (2.2)                     | 3.0 (1.3)                | 3.7 (1.4)               |
| Do you care about your partner?  | 4.3 (2.6)                     | 3.2 (1.3)                | 2.3 (1.8)               |
| <i>M</i>   | 4.7 (2.2)                     | 2.3 (1.6)                | 3.0 (1.6)               |
| <b>What did you think?</b>   |                               |                          |                         |
| There is a real human being in another room manipulating what is on the screen.          | 3.6 (2.6)                     | 3.0 (1.5)                | 3.2 (2.5)               |
| My partner's responses are completely programmed.  | 6.1 (1.2)                     | 5.2 (1.19)               | 5.8 (1.2)               |
| My partner has the ability to change responses and make decisions based on what you say. | 5.3 (1.6)                     | 5.0 (1.7)                | 4.5 (2.1)               |
| I didn't consider any of these possibilities.  | 5.9 (1.1)                     | 4.7 (0.8)                | 4.8 (1.9)               |

*Note.* Standard deviations are in parentheses. Participants followed the experimental procedure up to the first choice and then completed this questionnaire. Items (other than Question 1) were presented on Likert scales (1 = lowest value and 7 = highest value).

commitment (agreements to cooperate) and with cooperation, with the exception that participants' ratings of their competitiveness were correlated negatively with their total cooperation ( $r = -.93$ ). There were no differences among conditions in how well participants followed instructions over the six trials, whether they made a commitment, or whether they made extraneous comments to the experimenter or partner. All participants made choices on all trials.

Four items on the postchoice questionnaire checked whether participants across conditions similarly perceived the partner to be making real choices. There were no differences among conditions on the 7-point scale items "How much did your partner respond to you as a person?" and "To what degree were you concerned about the reactions of your partner as you made your investment decisions?" There also were no differences among conditions in which participants were asked if their partner seemed to be responding to their choices rather than "going it alone," or whether their partner's choices were "genuine" versus "fake or predetermined." These data suggest that participants in both the person and computer conditions were similarly treating the partner as making real choices and that experimenter demand did not differ among conditions of the study.

To assess attention to the partner, we asked participants to recall how many promises and suggestions their partners had made during the choice rounds. Participants in the person condition on average overestimated how many promises the partner had made (2.8 vs. the 2 actually made on Trials 1 and 4) and how many suggestions the partner had made (3.5 vs. the 2 actually made on Trials 2 and 5). Participants in the computer conditions generally were more accurate (face-voice = 2.4 promises and 2.2 suggestions; voice = 2.0 promises and 2.1 suggestions; text = 2.3 promises and 2.5 suggestions),  $F(3, 81) = 5.02$ ,  $p < .01$ . The contrast between the computer conditions and the person condition was significant,  $F(1, 81) = 13.75$ ,  $p < .001$ . These results suggest that participants paid closer attention to the computer partner than to the confederate (see Sproull et al., in press), but other explanations are possible too. Participants might have predicted or hoped for more commitment from the confederate than from the computer and adjusted their perceptions or their memories accordingly.

### *Participants' First Commitment and Choice*

Behavior on the first trial reveals participants' responses to a partner whose choice is unknown. We hypothesized that if people cooperate out of a feeling of group identity, participants would suggest cooperation and cooperate more with a person than with a computer. We proposed alternative hypotheses about cooperation with different computers. If simply having human-like features increases the likelihood of group identity, and if group identity is sufficient to generate cooperation, then cooperation might be greater, the more human-like the computer. However, if human-like features seem abnormal or provoke wariness, then people might reject and cooperate less with a computer partner that has more human-like features. Within the social contract theoretical framework, we hypothesized that participants would be more likely to trust and therefore to commit to cooperation with a person than with a computer but that they would cooperate more with a partner to whom they had

made a commitment than with one to whom they had not, even if their partner were a computer. Table 3 shows participants' commitments and choices on each trial.

On the first trial, the partner asked the participant to suggest a choice, and all but 3 participants did so. The partner always agreed with the participant's suggestion. A majority of the participants suggested cooperation, and there was no statistically significant difference among conditions. However, participants actually cooperated differentially across conditions,  $F(3, 82) = 4.0$ ,  $p = .01$ . Eighty percent of the participants in the person condition cooperated, approximately twice the 41% who cooperated in the computer conditions; the planned contrast of person versus computer conditions was significant,  $F(1, 82) = 9.9$ ,  $p < .01$ . Within the computer conditions, the percentage of cooperators was 32% in the computer face-voice condition, 43% in the computer voice condition, and 50% in the computer text condition. Ignoring commitment, this pattern suggests a trend against cooperation with more human-like computers, but the contrasts were not significant.

Figure 3 shows the extent to which participants kept their commitments. A repeated measures ANOVA of choice and commitment  $\times$  condition revealed a significant main effect of condition,  $F(3, 79) = 3.1$ ,  $p < .05$ , and a significant interaction effect,  $F(3, 79) = 3.1$ ,  $p < .05$ . The interaction effect indicates different trends for those who suggested competition and cooperation. Among the minority who proposed competition, 67% of those in the person condition and approximately 88% in the computer conditions kept their commitment; the contrasts are not significant ( $F < 1$ ). Among the majority who proposed cooperation, 94% in the person condition, 41% in the computer face-voice condition, 67% in the computer voice condition, and 90% in the computer text condition kept their commitments to cooperate. Here, the contrast of the person condition with the computer conditions was significant,  $F(1, 79) = 5.4$ ,  $p < .05$ ; the contrast of the computer face-voice condition and the other two computer conditions was significant,  $F(1, 79) = 6.0$ ,  $p < .05$ ; and the contrast of the computer voice with the computer text was not,  $F(1, 79) = 2$ .

### *Choices Across Trials*

Choices over all trials are shown in Figure 4. Over trials, cooperation increased slightly on Trial 2 and then declined,  $F(5, 410) = 10.1$ ,  $p < .001$ , and there was a strong main effect for condition,  $F(3, 82) = 10.86$ . The interaction term was not significant ( $F = 1.3$ ). The planned contrast between the person condition and the computer conditions was significant,  $F(1, 82) = 31.9$ ,  $p < .001$ , but the computer conditions did not differ. We next discuss comparisons across trials that address arguments about social identity and social contract explanations more specifically.

Choices on the first four trials, on which the amount and type of discussion were manipulated, but on which the partner cooperated consistently, allow for a preliminary evaluation of the role of discussion in differential responses to a human or computer partner. Table 2 and Figure 4 show choices across all trials. A repeated measures ANOVA across the first four trials revealed a strong main effect of condition,  $F(3, 82) = 11.4$ ,  $p < .001$ ; a main effect of trial,  $F(3, 246) = 8.3$ ,  $p < .001$ ; and no

**Table 3**  
*Number of Participants Who Made Commitments, Cooperated, and Kept Their Commitments With a Human or Computer Partner*

| Participants' response   | Experimental condition     |  |                                       |                                      |
|--|----------------------------|--|---------------------------------------|--------------------------------------|
|  | Person<br>( <i>n</i> = 20) | Computer<br>face-voice<br>( <i>n</i> = 25) | Computer<br>voice<br>( <i>n</i> = 21) | Computer<br>text<br>( <i>n</i> = 20) |
| Trial 1: Partner asked participant to propose (and cooperated) |                            |  |                                       |                                      |
| Proposed cooperation   | 16                         | 17   | 12                                    | 10                                   |
| Cooperated   | 15                         | 7  | 8                                     | 9                                    |
| Competed   | 1                          | 10   | 4                                     | 1                                    |
| % kept commitment  | 94%                        | 41%  | 67%                                   | 90%                                  |
| Proposed competition   | 3                          | 8  | 8                                     | 9                                    |
| Cooperated   | 1                          | 1  | 1                                     | 1                                    |
| Competed   | 2                          | 7  | 7                                     | 8                                    |
| % kept commitment  | 67%                        | 88%  | 88%                                   | 89%                                  |
| Did not propose  | 1                          | 0  | 1                                     | 1                                    |
| Cooperated   | 0                          | na   | 0                                     | 0                                    |
| Competed   | 1                          | na   | 1                                     | 1                                    |
| Total cooperation  | 16                         | 8  | 9                                     | 10                                   |
| % of participants who cooperated                               | 80%                        | 32%  | 43%                                   | 50%                                  |
| Trial 2: Partner proposed cooperation (and cooperated)         |                            |  |                                       |                                      |
| Proposed cooperation   | 20                         | 18   | 16                                    | 14                                   |
| Cooperated   | 19                         | 9  | 10                                    | 7                                    |
| Competed   | 1                          | 9  | 6                                     | 7                                    |
| % kept commitment  | 95%                        | 50%  | 63%                                   | 50%                                  |
| Proposed competition   | 0                          | 6  | 5                                     | 5                                    |
| Cooperated   | na                         | 1  | 0                                     | 0                                    |
| Competed   | na                         | 5  | 5                                     | 5                                    |
| % kept commitment  | na                         | 83%  | 100%                                  | 100%                                 |
| Did not propose  | 0                          | 1  | 0                                     | 1                                    |
| Cooperated   | na                         | 1  | na                                    | 1                                    |
| Competed   | na                         | 0  | na                                    | 0                                    |
| Total cooperation  | 19                         | 11   | 10                                    | 8                                    |
| % of participants who cooperated                               | 95%                        | 44%  | 48%                                   | 40%                                  |
| Trial 3: No discussion (partner cooperated)                    |                            |  |                                       |                                      |
| Total cooperation  | 9                          | 8  | 3                                     | 4                                    |
| % of participants who cooperated                               | 45%                        | 32%  | 14%                                   | 20%                                  |
| Trial 4: Partner asked participant to propose (and competed)   |                            |  |                                       |                                      |
| Proposed cooperation   | 10                         | 11   | 10                                    | 10                                   |
| Cooperated   | 9                          | 3  | 4                                     | 5                                    |
| Competed   | 1                          | 8  | 6                                     | 5                                    |
| % kept commitment  | 90%                        | 27%  | 40%                                   | 50%                                  |
| Proposed competition   | 5                          | 12   | 10                                    | 10                                   |
| Cooperated   | 0                          | 2  | 1                                     | 1                                    |
| Competed   | 5                          | 10   | 9                                     | 9                                    |
| % kept commitment  | 100%                       | 83%  | 90%                                   | 90%                                  |
| Did not propose  | 5                          | 2  | 1                                     | 0                                    |
| Cooperated   | 3                          | 0  | 0                                     | na                                   |
| Competed   | 2                          | 2  | 1                                     | na                                   |
| Total cooperation  | 12                         | 5  | 5                                     | 6                                    |
| % of participants who cooperated                               | 60%                        | 20%  | 24%                                   | 30%                                  |

(table continues)

Table 3 (continued)

| Participants' response                                 | Experimental condition     |  |                                       |                                      |
|--|----------------------------|--|---------------------------------------|--------------------------------------|
|  | Person<br>( <i>n</i> = 20) | Computer<br>face-voice<br>( <i>n</i> = 25) | Computer<br>voice<br>( <i>n</i> = 21) | Computer<br>text<br>( <i>n</i> = 20) |
| Trial 5: Partner proposed cooperation (and cooperated) |                            |  |                                       |                                      |
| Proposed cooperation                                   | 16                         | 15   | 13                                    | 13                                   |
| Cooperated   | 11                         | 7  | 4                                     | 3                                    |
| Competed   | 5                          | 8  | 9                                     | 10                                   |
| % kept commitment                                      | 65%                        | 47%  | 31%                                   | 23%                                  |
| Proposed competition                                   | 2                          | 10   | 7                                     | 6                                    |
| Cooperated   | 0                          | 0  | 0                                     | 0                                    |
| Competed   | 2                          | 10   | 7                                     | 6                                    |
| % kept commitment                                      | 100%                       | 100%                                       | 100%                                  | 100%                                 |
| Did not propose  | 1                          | 0  | 1                                     | 1                                    |
| Cooperated   | 0                          | na   | 1                                     | 0                                    |
| Competed   | 1                          | na   | 0                                     | 1                                    |
| Total cooperation                                      | 11                         | 7  | 5                                     | 3                                    |
| % of participants who cooperated                       | 55%                        | 28%  | 24%                                   | 15%                                  |
| Trial 6: No discussion (partner competed)              |                            |  |                                       |                                      |
| Total cooperation                                      | 4                          | 2  | 4                                     | 6                                    |
| % of participants who cooperated                       | 20%                        | 8%   | 14%                                   | 30%                                  |

significant interaction. The planned contrasts indicate that, on average, participants in the person condition cooperated more than participants in the computer conditions,  $F(1, 82) = 33.1$ ,  $p < .001$ , but the face-voice computer condition was not different from the other two computer conditions. Hence the trend for cooperation to decline with a more human-like interface was not repeated on the later trials. Possibly, participants in the computer conditions had by Trial 3 concluded that their computer partner would be reliably cooperative, making defection too tempting to ignore.

The strong impact of discussion and commitment, regardless of who—or what—was the discussion partner, is suggested by the sharp drop in cooperation in all conditions on Trial 3, a no-discussion trial. A contrast between Trials 1, 2, and 4, on which participants were induced to make an agreement, versus Trial 3, which discouraged any discussion, is highly significant,  $F(1, 246) = 12.0$ ,  $p < .001$ . Also, there was a (statistically insignificant) tendency for cooperation to increase from Trial 1 to Trial 2 (50% to 56%). If cooperation on the first two trials by the participant reinforced participants' confidence in cooperation, one might expect cooperation to have been sustained on Trial 3 even without discussion, but it was not. Twenty-eight participants cooperated on both Trials 1 and 2; fewer than half of those early cooperators cooperated on Trial 3, and the proportion was not significantly higher than it was among those who cooperated on only one of the first two trials. Moreover, on Trial 4, when participants once again had an opportunity to communicate, the proportion of participants who chose cooperation rose again in all conditions except in the computer face-voice condition. As on Trial 1, when the partner's communication was similar, there was no difference among conditions in the proportion of participants who made a commitment to cooperate, but there were differences in keeping commitments in the same direction as in Trial 1; that is, those in the face-voice condition

were more likely to break their promise,  $F(3, 82) = 4.5$ ,  $p < .01$ . Overall, the data for behavior over trials give support to social contract explanations of cooperation. Participants were much more likely to cooperate if they made an agreement and, as hypothesized, this pattern occurred whether their partner was a person or a computer. A counterbalanced series of trials would be required to test the explanation fully.

We examined choice over trials for evidence of strategic behavior. One strategy would be tit-for-tat matching of the partner's choices, but there was no evidence of this. Participants were not more likely to compete after the partner competed than after the partner cooperated. Another strategy might be to build a reputation for cooperation and then defect just before the partner does (Andreoni & Miller, 1993). The trend to compete increasingly over trials is consistent with this strategy, as is earlier defection in the computer conditions. Participants in the computer conditions might have thought their computer partner gullible, leading them to defect early and to use promises as a ruse.

One way to evaluate promises as strategic messages is to see what happened once the partner defected on Trial 4. This behavior would be expected to encourage all participants to use an agreement to cooperate strategically on Trial 5. Because the partner's verbal inducements were approximately the same on Trials 2 and 5, by comparing how participants kept their agreements to cooperate on these trials we can estimate the impact of the partner's deception on participants' strategic responses. Although an analysis of only those participants who agreed to cooperate indicates that on Trial 2 they were more likely to keep their agreement (66%) than were participants who agreed to cooperate on Trial 5 (43%;  $F[3, 45] = 7.72$ ,  $p < .001$ ), all behavior became more negative on Trial 5. An internal analyses of condition  $\times$  trial (2 vs. 5)  $\times$  commitment

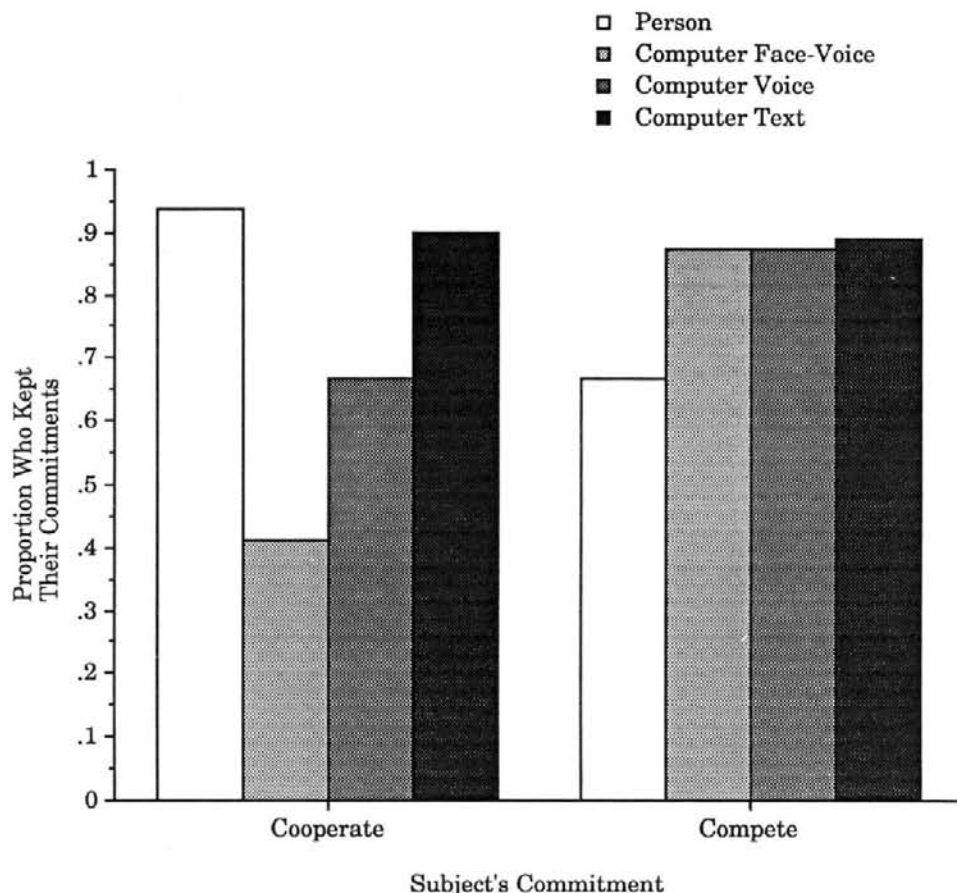


Figure 3. Proportion of participants who kept their commitments on Trial 1.

(cooperate or compete)  $\times$  kept agreement (yes or no), evidenced no interaction effects at all; interactions of trial with agreements to cooperate would be expected if on Trial 5 participants were less likely to keep an agreement to cooperate. Though here we see little evidence of strategic behavior, this might be due to self-selection of participants preferring cooperation as the trials proceeded.

*Postchoice evaluations of the partner.* Measures of feelings of partnership and evaluations of the partner after the choice trials are presented in Table 4 and summarized in Figure 5.<sup>6</sup> Five items measured feelings of partnership. There was a significant main effect for condition,  $F(3, 79) = 3.3, p < .05$ ; and for item,  $F(4, 316) = 21.1$ ; but no interaction. The planned contrast comparing the person condition with the computer conditions was highly significant,  $F(1, 79) = 7.2, p < .01$ , indicating stronger feelings of partnership in the person condition. Within the computer conditions, the difference between the computer face-voice condition and the other computer conditions was marginally significant, indicating participants had stronger feelings of partnership in the face-voice condition,  $F(1, 79) = 2.6, p = .10$ . The voice and text computer conditions did not differ.

Social evaluation differed significantly among conditions,  $F(3, 81) = 12, p < .001$ , and there were significant item differ-

ences and a Condition  $\times$  Item interaction,  $F(12, 324) = 2.9, p < .01$ . On inspection, these effects reflect higher social evaluation of the confederate than the computer partner; the planned contrast comparing the person condition with the computer conditions was significant,  $F(1, 81) = 14.3, p < .01$ . The effects also reflect low social evaluation ratings of the partner in the computer face-voice condition, including exceptionally low ratings of the attractiveness, cheerfulness, and warmth of the computer face-voice partner. The planned contrast comparing the computer face-voice condition with the other computer conditions was significant,  $F(1, 81) = 18.1, p < .01$ . The voice and text computers did not differ.

Like ratings of partnership, intelligence evaluation was highest in the person condition and next highest in the computer face-voice condition,  $F(3, 82) = 5.5, p < .01$ . There were no significant item differences and no interaction of item with condition. The planned contrast of the person condition with the computer conditions was significant,  $F(1, 82) = 16.0, p < .01$ , but the computer conditions did not differ. These findings rep-

<sup>6</sup> Some ratings were presented to participants in a nonintuitive direction (more positive = lower score). We have reversed these scores so that all ratings are reported with a higher score equivalent to a more positive valence.

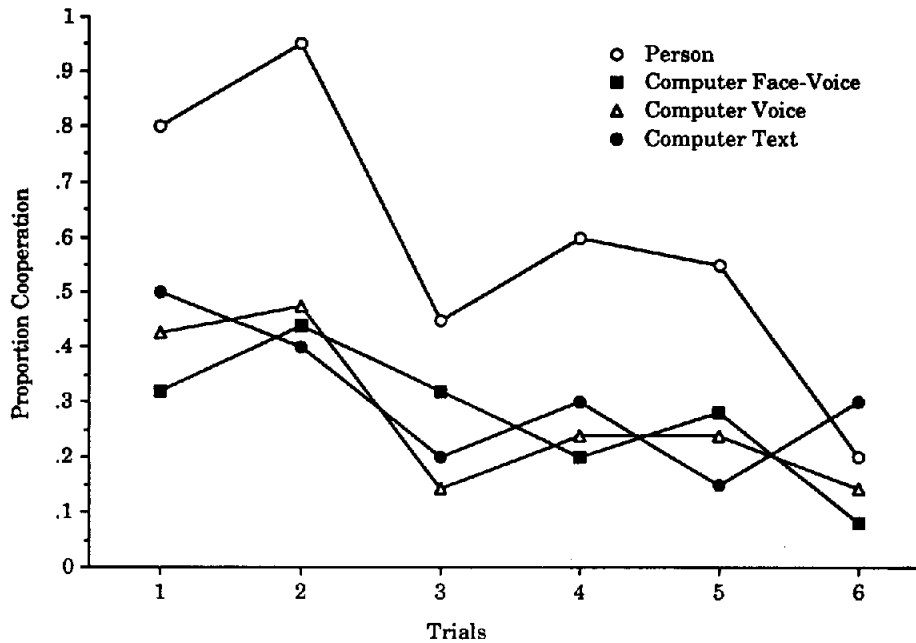


Figure 4. Cooperation over trials.

licate those of a previous study by Sproull et al. (in press), in which participants gave lower social evaluations but not lower intelligence evaluations to a computer face-voice interviewer than to a computer text interviewer.

Other evaluations generally did not differ. As in our previous work, Warner-Sugarman potency ratings were not different across conditions, nor were ratings of how similar participants felt to the partner. Participants in the person and computer text condition said the partner resembled someone they knew, as compared with the partner in the computer face-voice and computer voice conditions,  $F(3, 82) = 5.8, p < .01$ , but the contrasts were not significant.

#### Effects of Other Variables

There were no effects of gender, age, native language, or computer expertise on commitments, choices, or whether participants kept their commitments. We replicated a previous finding (Sproull et al., in press) that women gave less positive social evaluations to the face-voice computer than men did. In this study an interaction of condition and gender,  $F(3, 8) = 2.9, p < .05$ , reflected the case that women gave higher ratings to the human and computer text partner than men did but gave lower ratings to the face-voice and voice computer partners.

#### Discussion

Our results suggest that people can respond socially to technology, that some of the processes influencing relations with a person also affect people's relations with a computer, and that people follow social rules in interacting with a computer. Cooperation with a computer partner in a dilemma was half that with a person, but cooperation after discussion with a computer

partner was still substantial—higher than with uncommunicative computers and much higher when comparing the best-liked computer partner (computer text condition = 50%) with computer partners of the past (e.g., < 35% in Abrie & Kahan [1972]; 20–30% in Shure & Meeker [1968]).

The results lend support to both social identity and social contract explanations of cooperation following discussion and suggest that identity or commitment processes might affect interactions with a computer. On Trial 1, which was uncontaminated by the partner's choice, a majority of participants across conditions suggested cooperation, but they were more likely to break their promises to cooperate when the partner was a computer, especially when the computer was more human-like. They also rated the human-like computer as the least friendly, cheerful, attractive, and warm of all the partners. One interpretation of these findings is that participants simply rejected the human-like computer as a partner and then pursued a self-interested strategy. This interpretation is consistent with research on stigma and novelty (e.g., Fiske, 1980; Langer et al., 1976). However, participants' ratings of their feelings of partnership with the human-like computer were only slightly lower than their feelings of partnership toward the confederate. We suggest another interpretation of the results of Trial 1, which is more in keeping with the ratings of partnership, indicating minimal identity with the computer partners, especially with the human-like partner. Participants might have formed a group identity of sorts with their computer partner but disliked and took advantage of it, especially when it was human-like but imperfect. Hogg (1992) argued that intragroup downgrading of marginal group members occurs when there is intergroup competition and a group member fails to meet expectations on important group dimensions (pp. 118–125). Marques (1990)

Table 4  
Mean Postchoice Evaluations of the Partner

| Measures   | Partner manipulation       |  |                                       |                                      |
|--|----------------------------|--|---------------------------------------|--------------------------------------|
|  | Person<br>( <i>n</i> = 20) | Computer<br>face-voice<br>( <i>n</i> = 25) | Computer<br>voice<br>( <i>n</i> = 21) | Computer<br>text<br>( <i>n</i> = 20) |
| <b>Feelings of partnership</b>   |                            |  |                                       |                                      |
| How strong a feeling of being part of a partnership did you have?  | 3.8 (1.7)                  | 3.0 (1.1)                                  | 2.7 (1.9)                             | 2.2 (1.6)                            |
| If you had the chance would you rather [not] have had a different partner?   | 4.1 (1.7)                  | 4.4 (1.6)                                  | 3.7 (2.0)                             | 3.6 (2.2)                            |
| If you had the chance would you like to work together with your same partner again?                                    | 4.2 (1.6)                  | 4.9 (1.7)                                  | 3.9 (1.6)                             | 4.4 (1.7)                            |
| To what degree did you feel like you were playing the investment game for yourself versus playing for the partnership? | 2.9 (2.0)                  | 2.6 (2.0)                                  | 2.1 (1.6)                             | 2.0 (1.5)                            |
| How glad would you feel if your partner won the lottery exchanging investment credits for dollars?                     | 5.4 (1.6)                  | 3.6 (2.2)                                  | 4.0 (1.8)                             | 3.9 (1.3)                            |
| $\bar{X}$  | 4.1 (1.9)                  | 3.7 (2.0)                                  | 3.4 (1.9)                             | 3.2 (1.9)                            |
| <b>Social evaluation scale (Warner-Sugarman)</b>   |                            |  |                                       |                                      |
| Attractive   | 4.2 (1.4)                  | 2.6 (1.7)                                  | 3.9 (1.7)                             | 3.7 (1.3)                            |
| Cheerful   | 5.3 (1.8)                  | 3.0 (1.2)                                  | 4.5 (1.3)                             | 5.4 (1.2)                            |
| Friendly   | 5.8 (1.7)                  | 4.4 (1.3)                                  | 5.2 (1.3)                             | 5.5 (1.6)                            |
| Optimistic   | 5.7 (1.2)                  | 4.9 (1.3)                                  | 4.5 (1.0)                             | 5.0 (1.2)                            |
| Warm   | 5.1 (1.4)                  | 3.0 (1.2)                                  | 4.0 (1.2)                             | 4.5 (1.2)                            |
| $\bar{X}$  | 5.3 (1.6)                  | 3.6 (1.6)                                  | 4.4 (1.4)                             | 4.8 (1.4)                            |
| <b>Intelligence evaluation scale (Warner-Sugarman)</b>   |                            |  |                                       |                                      |
| Knowledgeable  | 5.5 (1.1)                  | 4.5 (1.5)                                  | 4.4 (1.5)                             | 4.2 (1.8)                            |
| Competent  | 5.7 (1.1)                  | 4.4 (1.6)                                  | 4.2 (1.3)                             | 4.5 (1.6)                            |
| Responsible  | 5.2 (1.5)                  | 4.8 (1.4)                                  | 4.8 (1.2)                             | 4.4 (1.5)                            |
| Sensible   | 5.4 (1.3)                  | 4.5 (1.5)                                  | 4.0 (1.3)                             | 4.5 (1.7)                            |
| Intelligent  | 5.6 (1.3)                  | 4.2 (1.6)                                  | 3.8 (1.4)                             | 3.9 (1.6)                            |
| $\bar{X}$  | 5.5 (1.3)                  | 4.5 (1.5)                                  | 4.3 (1.4)                             | 4.3 (1.6)                            |

Note. Standard deviations are in parentheses. All items were presented on 7-point Likert rating scales. Some items were presented in reversed order, but in this table a higher mean always represents a more positive evaluation.

called this phenomenon the *black sheep effect*. Participants who had a human-like computer partner met these conditions in that other partnerships were competing with them for prizes, and the partner differed on relevant dimensions. The other computers differed more from the participants in an absolute sense, but not on visual facial dimensions people use to evaluate others socially and to trust them (Sproull et al., in press). A better test of this explanation, of course, awaits future research.

Social contract arguments as to why people cooperate in dilemmas rest on the role of commitment during discussion. Our results show strong effects of commitments during discussion on cooperation. The majority of participants on Trial 1 made a commitment to cooperate during discussion and kept to this commitment, though they kept their promises less to a computer, especially to a human-like computer. Furthermore, across trials, cooperation was much higher when participants had a chance to make a commitment to cooperate, and participants who promised to cooperate were much more likely to cooperate than participants who did not, regardless of whether the partner was a computer or a person. Internal analyses of choices on each of the discussion trials (1, 2, 4, and 5) by participants who made a commitment to cooperate or compete show highly

significant effects of commitment on choice on each of these trials: Trial 1,  $F(1, 75) = 31.1, p < .001$ ; Trial 2,  $F(1, 77) = 16.0, p < .001$ ; Trial 4,  $F(1, 70) = 20.4, p < .001$ ; Trial 5,  $F(1, 75) = 13.6, p < .001$ .

We derived two reasons from the literature as to why people keep their commitments—out of a sense of obligation to the partnership or group and out of a sense of obligation to the self. For those persons feeling obligated to the partnership, the nature of the partner might well affect the strength of that feeling. For example, would the partner appreciate and reciprocate honorable behavior? Would the partner be worthy of loyalty? For persons feeling obligated to themselves, the partner might make less of a difference—to some people a promise is a promise. We cannot glean from our data whether one or both of these processes was operating in this study, but our results indicate that participants did not value commitment above all other considerations. Participants who made commitments to a computer, especially to a human-like computer, seem to have interpreted their commitments differently than did participants who made commitments to a person. This kind of differential commitment in a dilemma seems worthy of future exploration.

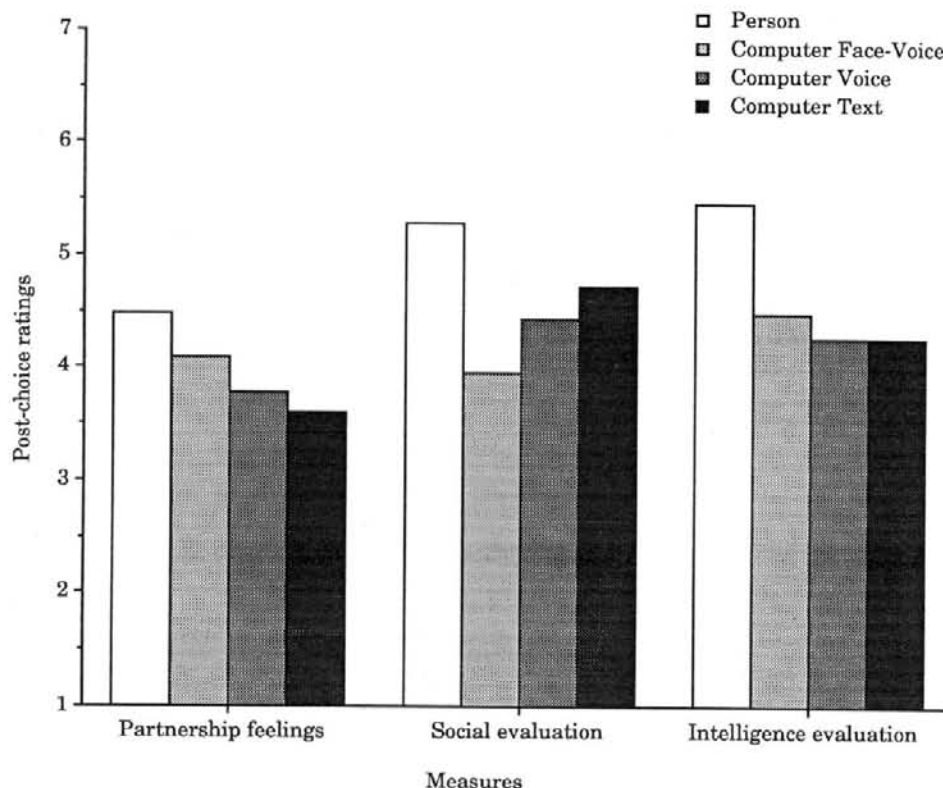


Figure 5. Postchoice feelings of partnership, sociability evaluation of partner, and intelligence evaluation of partner.

### Limitations

Our results, and our interpretation of them in the context of social identity and social contract approaches, are constrained by the design, procedure, and technology we used and require more empirical and theoretical development. For instance, we gave social identity and social contract processes a boost by referring to the computer and to the participant as partners and by encouraging pregame conversation. Also, by failing to counterbalance discussion and no-discussion trials, we cannot rule out the possibility that having discussion first might have increased its apparent effect on cooperation. Many reasons other than the ones we proposed might be advanced to explain why participants would cooperate with a computer: to please or to avoid displeasing the experimenter, because it is a dominant response given by participants when they are anxious; because cooperation makes the computer seem less frightening, because the computer seems guided by a real person, or because the experimenter treats the computer as though it were human.<sup>7</sup> We cannot positively rule out these explanations, but questionnaire ratings of the experimenter's behavior, of the participants' feelings and emotions, and of participants' perceptions of the autonomy of the partner did not differ across conditions and do not support these conjectures. Moreover, these explanations do not account for participants' differential responses to the human-like computer as compared with the other computers.

The somewhat negative reactions to the talking face com-

puter in this study might have resulted from idiosyncrasies in the particular computer interface we used. A more perfect representation of a face and voice in a computer display might increase rather than decrease cooperative behavior. A more perfect representation might increase the computer partner's credibility or remind people more of cooperation schemas. Recent research suggests that visual fidelity in computer representations increases involvement and learning (Christel, 1994), but the reasons for these effects are not well understood. A realistic synthetic partner for mixed-motive paradigms awaits future technology improvements in both speech and person representation. Until then, experimenters have a variety of options: video displays, animation, and the better talking faces that continue to come out of computer science laboratories.

Our discussion of the imperfect talking face computer assumes that realism is a key variable in communication and, if so, that as it becomes possible to increase the realism of the representation of a person, cooperation should increase. Alternatively, consider a person interacting with a computer representation of Kermit the Frog or Minnie Mouse. Would a charming cartoon character that is clearly not human prompt more cooperation? A positive answer would suggest that likability or resemblance to young children or pets rather than realism is the key element in eliciting communication that leads to social

<sup>7</sup> This list was suggested by an anonymous reviewer.



identity and social contract. If an attribute such as the face-to-body ratio is important, then ever more realistic computers might not increase cooperation. They would not necessarily be attractive, and even minor deviations from reality would remind people of the fundamentally unreal nature of the machine. If we are to better understand the fundamental properties of social identification and social contract, we need to understand what aspects of an artificial partner matter in producing social responses.

### Conclusion

This research, though exploratory, represents a new perspective on what it means to be social. Human social behavior, it seems, does not depend on interaction with people or even categorization in a human group. Social psychological research dating from the 1930s on social facilitation and the development of strong ties and in-group identification has emphasized the impact on social behavior of the presence of people and interaction among them. For example, Lewin's (1943, 1948) studies of leadership and group discussion and Sherif et al.'s (1961) studies of conflict and cooperation in a boys' camp showed how direct contact in a group and reciprocity encouraged cooperation. To explain cooperation with a computer, we need a theory of cooperation that can arise in the absence of people.

Theorists have long been interested in cooperation when people cannot communicate directly with their beneficiaries (Mauss 1925/1967). Macy (1990) argued that group members can learn to cooperate or contribute to group welfare in the absence of direct communication with others if others' cooperative behavior is visible or reinforced. He argued that when people observe others contributing and see rewards to the public good they may develop a norm of cooperation. Once a critical mass of contributors is achieved (e.g., Marwell, Oliver, & Prahl, 1988), the legitimacy and value of cooperation can be self-reinforcing. In a related argument, Cialdini and his colleagues (Cialdini, Reno, & Kallgren, 1990) argued that the visibility of others' cooperative behavior or cooperative outcomes reminds people that others cooperate in the situation, increasing the salience of a norm of cooperation. The residue of cooperative or antisocial behavior, such as litter on a street, can trigger normative responses. Generally, this work suggests that social behavior—positive or negative—is triggered by situational cues, which need not be other people.

Our work suggests that commitments can be elicited fairly easily, even by a machine. If keeping their word is important to people's self-identity and feelings of self-worth, and if they gain personal and social benefits from sticking to social contracts, then they might value even commitments to a computer. Cooperation to keep a commitment, in this view, is social insofar as honoring commitments is learned in society; the consequences of cooperation and defection are, of course, social, but social interaction with people or knowledge of specific human beneficiaries need not be presumed. This argument suggests that people keep commitments, whether to a person, group, dog, or computer, not because they believe each of these is human, but because they, themselves, are human.

Critics have long complained that research in artificial intelligence, which exploits parallels between computer behavior and

human behavior, is deeply demeaning to the essential humanness of people. Herbert Simon, in response to those critics, has said

I think that some people's anxiety about the computer is that it will relieve us of another source of our uniqueness as the only kind of system in this universe that can think. To say that computers can do what we do when we think and use language and discover is to challenge another area of human uniqueness . . . Why don't we try to get our sense of worth in this world, not in uniqueness, not in being apart from the rest of nature, but in being a general and integral part of nature. (Simon, 1987, pp. 14–15)

In the spirit of that view, we suggest that by understanding how people interact with technology that has more or fewer human characteristics, we not only can understand more about how people interact with technology but also understand more about what makes us social.

### References

- Abric, J. C., & Kahan, J. (1972). The effects of representations and behavior in experimental games. *European Journal of Social Psychology, 2*, 129–144.
- Allen, K. M., Blascovich, J., Tomaka, J., & Kelsey, R. M. (1991). Presence of human friends and pet dogs as moderators of autonomic responses to stress in women. *Journal of Personality and Social Psychology, 61*, 582–589.
- Amabile, T. M. (1983). Brilliant but cruel: Perceptions of negative evaluators. *Journal of Personality and Social Psychology, 19*, 146–156.
- Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal, 103*, 570–585.
- Apfelbaum, E. (1974). On conflicts and bargaining. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 103–156). New York: Academic Press.
- Ball, E., Ling, D. T., Pugh, D., Skelly, T., Stankosky, A., & Thiel, D. (1994, April). *ReActor: A system for real-time, reactive animations*. Demonstration at the ACM Conference on Human Factors in Computing Systems CHI '94, Boston.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*, 243–267.
- Belk, R. W. (1988). Possessions and the extended self. *Journal of Consumer Research, 15*, 139–168.
- Binick, Y. M., Westbury, C. F., & Servan-Schreiber, D. (1989). Case histories and shorter communications. *Behavioral Research Therapy, 27*, 303–306.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307–324.
- Burnstein, E. (1969). Interdependence in groups. In J. Mills (Ed.), *Experimental social psychology* (pp. 307–405). Toronto: MacMillan.
- Caporael, L. R., Dawes, R. M., Orbell, J. M., & van de Kragt, A. J. C. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences, 12*, 683–699.
- Chen, X., & Komorita, S. S. (1994). The effects of communication and commitment in a public goods social dilemma. *Organizational Behavior and Human Decision Processes, 60*, 367–386.
- Christel, M. G. (1994). The role of visual fidelity in computer-based instruction. *Human-Computer Interaction, 9*, 183–223.
- Cialdini, R. B. (1984). *Influence: The new psychology of modern persuasion*. New York: Quill.
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory

- of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 1015-1026.
- Csikszentmihalyi, M., & Rochberg-Halton, E. (1981). *The meaning of things: Domestic symbols and the self*. Cambridge, England: Cambridge University Press.
- Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, 35, 1-11.
- Dawes, R. M., Orbell, J., & van de Kragt, A. J. (1988). Not me or thee but we: The importance of group identity in eliciting cooperation in dilemma situations. *Acta Psychologica*, 68, 83-97.
- Deane, P. D. (1993). On metaphoric inversion. *Metaphor and Symbolic Activity*, 8, 111-126.
- Dennett, D. C. (1988). Precis of the intentional stance. *Behavioral and Brain Science*, 11, 495-546.
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, 2, 265-279.
- Dittmar, H. (1992). *The social psychology of material possessions: To have is to be*. New York: St. Martin's Press.
- Ekman, P. (Ed.). (1982). *Emotion in the human face* (2nd ed.). Cambridge, England: Cambridge University Press.
- Elliott, C. (1994, July). Research problems in the use of a shallow artificial intelligence model of personality and emotion. *Proceedings of the Twelfth National Conference on Artificial Intelligence* (pp. 9-15), Seattle, Washington.
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889-906.
- Folkes, V. S., & Sears, D. O. (1977). Does everybody like a liker? *Journal of Experimental Social Psychology*, 13, 505-519.
- Friedman, B. (1995, May). *It's the computer's fault: Reasoning about computers as moral agents*. Paper presented at the ACM Conference on Human Factors in Computing Systems CHI '95, Denver, Colorado.
- Friedmann, E., Katcher, A. H., Lynch, J. J., & Thomas, S. A. (1980). Animal companions and one-year survival of patients after discharge from a coronary care unit. *Public Health Reports*, 95, 307-312.
- Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2, 167-177.
- Gozzi, R., Jr. (1989). Metaphors that undermine human identity. *Contents*, 46, 49-53.
- Hayes-Roth, B., Sincoff, E., Brownston, L., Huard, R., & Lent, B. (1995, May). *Directed improvisation with animated puppets*. Demonstration at the ACM Conference on Human Factors in Computing Systems CHI '95, Denver, Colorado.
- Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. Hertfordshire, England: Harvester Wheatsheaf.
- Hogg, M. A., & McGarty, C. (1990). Self-categorization and social identity. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 10-27). New York: Springer-Verlag.
- Insko, C. A., Hoyle, R. H., Pinkley, R. L., Hong, G., Slim, R., Dalton, B., Lin, Y., Ruffin, P. P., Dardis, G. J., Bernthal, P. R., & Schopler, J. (1988). Individual-group discontinuity: The role of a consensus rule. *Journal of Experimental Social Psychology*, 24, 505-519.
- Ito, K. S., Hayamizu, S., & Tanaka, H. (1992). Continuous speech recognition by context-dependent phonetic HMM and an efficient algorithm for finding N-best sentence hypotheses. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 121-124).
- Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.
- Kerr, N. L., & Kaufman-Gilliland, C. M. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology*, 66, 513-529.
- Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief*. New York: Academic Press.
- Kleck, R. E. (1968). Psychological stigma and nonverbal cues emitted in face-to-face interaction. *Human Relations*, 21, 19-28.
- Kleck, R., Buck, P. L., Goller, W. L., London, R. S., Pfeiffer, J. R., & Vučković, D. P. (1968). The effect of stigmatizing conditions on the use of personal space. *Psychological Reports*, 23, 111-118.
- Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, 46, 183-207.
- Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, 46, 1044-1057.
- Langer, E. J., Taylor, S. E., Fiske, S., & Chanowitz, B. (1976). Stigma, staring and discomfort: A novel-stimulus hypothesis. *Journal of Experimental Social Psychology*, 12, 451-463.
- Laurel, B. (1990). Interface agents: Metaphors with character. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 355-365). New York: Addison-Wesley.
- Lewin, K. (1943). Forces behind food habits and methods of change. *Bulletin of the National Research Council*, 108, 35-65.
- Lewin, K. (1948). *Resolving social conflicts*. New York: Harper.
- Macy, M. W. (1990). Learning theory and the logic of critical mass. *American Sociological Review*, 55, 809-826.
- Marques, J. M. (1990). The black-sheep effect: Out-group homogeneity in social comparison settings. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 131-151). Hemel Hempstead, England: Harvester Wheatsheaf.
- Marwell, G. P., & Ames, R. E. (1981). Economists free ride. Does anyone else? *Journal of Public Economics*, 15, 295-310.
- Marwell, G., Oliver, P., & Prahl, R. (1988). Social networks and collective actions: A theory of critical mass III. *American Journal of Sociology*, 94, 502-532.
- Mauss, M. (1967). *The gift*. New York: Norton. (Original work published 1925)
- Nagao, K., & Takeuchi, A. (1994, April). *Social interaction: Multimodal conversation with social agents*. (Technical paper No. SCSL-TR-94-005). Tokyo: Sony Computer Science Laboratory.
- Nasar, S. (1994, February 27). *The American economy, Back on top*. The New York Times, p. 6.
- Nass, C. I., & Steuer, J. (1993). Computers, voices, and sources of evaluation. *Human Communication Research*, 19, 504-527.
- Nass, C. I., Steuer, J., Henriksen, L., & Dryer, D. C. (1994). Machines and social attributions: Performance assessments of computers subsequent to "self-" or "other-" evaluations. *International Journal of Man-Machine Studies*, 40, 543-559.
- Nemeth, C. (1972). A critical analysis of research utilizing the prisoner's dilemma paradigm for the study of bargaining. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 203-234). New York: Academic Press.
- Orbell, J. M., Dawes, R. M., & van de Kragt, A. J. C. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology*, 54, 811-819.
- Orbell, J. M., van de Kragt, A. J., & Dawes, R. M. (1991). Covenants without the sword: The role of promises in social dilemma circumstances. In K. J. Koford & J. B. Miller (Eds.), *Social norms and economic institutions* (pp. 117-148). Ann Arbor: University of Michigan Press.
- Oskamp, S. (1971). Effects of programmed strategies on cooperation

- in the prisoner's dilemma and other mixed-motive games. *Journal of Conflict Resolution*, 15, 225-259.
- Ostrom, E., Walker, J. M., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, 86, 404-417.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363-392.
- Reeves, B., Lombard, M., & Melwani, G. (1992). *Faces on the screen: Pictures or natural experience?* Unpublished manuscript, Stanford University.
- Resnick, P. V., & Lammers, H. B. (1985). The influence of self-esteem on cognitive responses to machine-like versus human-like computer feedback. *Journal of Social Psychology*, 125, 761-769.
- Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, 7, 58-92.
- Schlenker, B. R. (Ed.). (1985). *The self and social life*. New York: McGraw-Hill.
- Schneiderman, B. (1987). *Designing the user interface: Strategies for effective human-computer interaction*. Boston: Addison-Wesley.
- Schopler, J., Insko, C. A., Graetz, K. A., Drigotas, S., Smith, V. A., & Dahl, K. (1993). Individual-group discontinuity: Further evidence for mediation by fear and greed. *Personality and Social Psychology Bulletin*, 19, 419-431.
- Searle, J. R. (1981). Minds, brains, and programs. In D. R. Hofstadter & D. C. Dennett (Eds.), *The mind's I* (pp. 353-372). Toronto: Bantam.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers' Cave experiment*. Norman, OK: University Book Exchange.
- Shamir, B. (1991). Meaning, self and motivation in organizations. *Organization Science*, 12, 405-424.
- Shure, G., & Meeker, R. J. (1968). Empirical demonstration of normative behavior in the prisoner's dilemma. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 3, 61-62.
- Simon, H. (1987). Computers and society. In S. Kiesler & L. Sproull (Eds.), *Computing and change on campus* (pp. 4-15). Cambridge, England: Cambridge University Press.
- Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, 32, 1492-1512.
- Sproull, L., Walker, J., Subramani, R., Kiesler, S., & Waters, K. (in press). When the interface is a face. *Human-Computer Interaction*.
- Tajfel, H. (1959). Quantitative judgment in social perception. *British Journal of Psychology*, 50, 16-29.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.
- Takemura, H., & Kishino, F. (1992). Cooperative work environment using virtual workspace. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, New York (pp. 226-232).
- Tedeschi, J. T. (1981). *Impression management theory and social psychological research*. New York: Academic Press.
- Toothaker, L. E. (1993). *Multiple comparison procedures*. Newbury Park, CA: Sage.
- Turkle, S. (1984). *The second self: Computers and the human spirit*. New York: Simon & Schuster.
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15-36). Cambridge, England: Cambridge University Press.
- Voissem, N. H., & Sistrunk, F. (1971). Communication schedule and cooperative game behavior. *Journal of Personality and Social Psychology*, 19, 160-167.
- Warner, R. M., & Sugarman, D. B. (1986). Attributions of personality based on physical appearance, speech, and handwriting. *Journal of Personality and Social Psychology*, 50, 792-799.
- Waters, K. (1987). A muscle model for animating three-dimensional facial expressions. *Computer Graphics*, 21, 17-24.
- Waters, K., & Levergood, T. M. (1995). DECface: A system for synthetic face applications. *Multimedia Tools and Applications*, 1, 1-16.

Received June 7, 1994

Revision received May 26, 1995

Accepted July 26, 1995 ■