



# Computers are Social Actors

Clifford Nass, Jonathan Steuer, and Ellen R. Tauber

Department of Communication  
Stanford University  
Stanford, CA 94305-2050, USA  
+1.415.723.5499

nass@leland.stanford.edu, jonathan@casa.stanford.edu, ellen@cs.stanford.edu

## ABSTRACT

This paper presents a new experimental paradigm for the study of human-computer interaction. Five experiments provide evidence that individuals' interactions with computers are fundamentally social. The studies show that social responses to computers are not the result of conscious beliefs that computers are human or human-like. Moreover, such behaviors do not result from users' ignorance or from psychological or social dysfunctions, nor from a belief that subjects are interacting with programmers. Rather, social responses to computers are commonplace and easy to generate. The results reported here present numerous and unprecedented hypotheses, unexpected implications for design, new approaches to usability testing, and direct methods for verification.

**KEYWORDS:** Anthropomorphism, Agents, Voice, Speech, Social Psychology, Gender, Design

## INTRODUCTION

What can we learn about human-computer interaction if we show that the human-computer relationship is fundamentally social? What can we predict and test if we assume that individuals are biased toward a social orientation; that when people sit down at a computer, they interact socially?

The present research provides a wide range of experimental evidence that a limited set of characteristics associated with humans provides sufficient cues to encourage users to exhibit behaviors and make attributions toward computers that are nonsensical when applied to computers but appropriate when directed at other humans. Thus, we demonstrate that users can be induced to elicit a wide range of social behaviors, even though users know that the machines do not actually possess feelings, "selves," genders, or human motivations.

The approach is as follows:

1. Pick a social science finding (theory and method) which concerns behavior or attitude toward humans. The studies presented here draw from social psychology and sociology.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

CHI94-4/94 Boston, Massachusetts USA

© 1994 ACM 0-89791-650-6/94/0072...\$3.50

2. Change "human" to "computer" in the statement of the theory.
3. Replace one or more humans with computers in the method of the study.
4. Provide the computer with characteristics associated with humans: (a) language output [1]; (b) responses based on multiple prior inputs [2]; (c) the filling of roles traditionally filled by humans [3]; and (d) the production of human-sounding voices [4,5,6,7].
5. Determine if the social rule still applies.

In this paper, we report successful application of our approach in five studies. The first study answers the question, "Will users apply politeness norms to computers?" The second study answers the question, "Will users apply the notions of 'self' and 'other' to computers?" The third study answers the question, "On what basis do users distinguish computers as 'self' or 'other' — the voice or the box?" The fourth study answers the question, "Will users apply gender stereotypes to computers?" Finally, the fifth study answers the question, "If people do respond socially to computers, is it because they feel that they are interacting with the computer or with some other agent, such as the programmer?" and, "Who or what do users think of when the a computer says 'I'?"

In sum, the basic question in the present studies, and a question that has not previously been answered, is, "Which social rules will people apply to computers?" A subsidiary question is how powerful are the rules; that is, can one easily generate these responses or do they only occur rarely?

A crucial point about this research is that all of these studies involve experienced computer users. Thus, none of the subjects' responses resulted merely from the novelty of using a computer, or from some misunderstanding or fallacious belief about the capabilities of computers. [8]

In the remainder of this paper, we outline the general method for the five experiments we performed. We then describe the specific methods and results for each experiment. Finally, we highlight theoretical and design implications for both the individual studies and for the experimental paradigm as a whole.



**GENERAL METHOD**

All five laboratory studies described below share the same basic experimental method. 180 computer-literate college students volunteered to participate in various experiments involving a computer tutor. Each experiment lasted approximately 45 minutes.

After entering the computer laboratory, the experimenter told each subject that he or she would prepare for a test with the assistance of a computerized tutoring session. Subjects used computers for three distinct sessions: tutoring, testing, and evaluation. After the evaluation session, subjects completed a questionnaire regarding their attitudes about the tutoring, testing, and evaluation sessions.

The computers used were NeXT workstations with 17" 2-bit greyscale monitors. The experimental applications were developed in-house using the NeXTStep 3.0 Application Kit and Interface Builder (Copies of the applications and questionnaires are available via anonymous ftp from srct.stanford.edu (36.125.0.127)). The interface was window-based; subjects interacted with the program by pressing on-screen buttons with a mouse. No use of menus or keyboards was required.

Each screen displayed a brief description of what information was being output (e.g., "Instructions," "Fact #7," "Evaluation 3") and one or more on-screen buttons which could be highlighted by clicking the mouse on the button. The button labeled "CONTINUE" indicated that the subject had heard the requisite information, had provided a response (if needed), and was ready to continue the tutoring, testing, or evaluation session. The button labeled "REPEAT" indicated that the subject wished to re-hear the information associated with the screen. A matrix of numbered buttons was used for the response scales for the tutoring and testing session. Visual characteristics of the interface (font, size, style, button and layout, etc.) were identical across conditions and across machines.

Sound was recorded to a computer hard disk in 16-bit monophonic format at a sampling rate of 44.1 kHz, using a

Singular Solutions A/D 64X analog-to-digital converter and a Sennheiser 421 dynamic microphone. Sound was played through a separate Fostex 3401 AVB amplified speaker concealed behind each computer monitor.

In all experiments using more than one voice, the complete set of sounds required for the experiment was recorded by at least three individuals. This allowed the particular voice used for each subject to be varied independently of the condition, thereby minimizing the possibility that the experimental effects resulted from the particular voices used, rather than from the experimental manipulations.

A summary of all five experiments is presented in Table 1; the basic experimental setup is depicted in Figure 1.

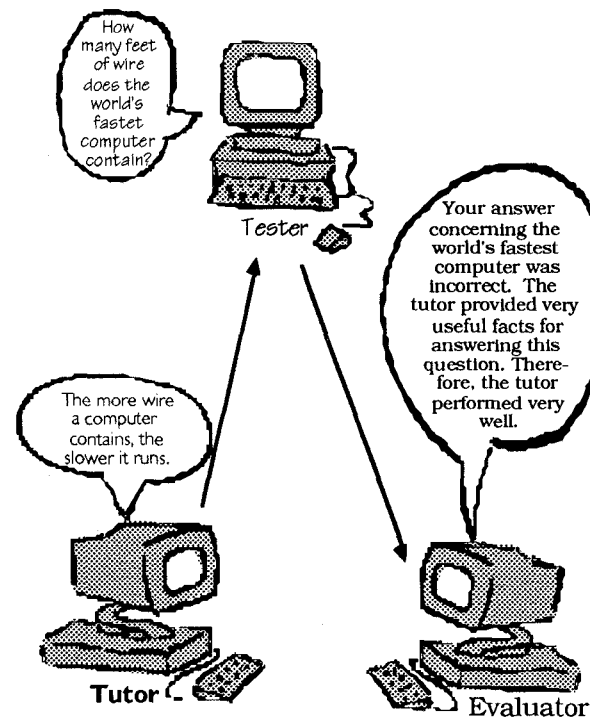


Figure 1: Overview of Lab Setup (Example: Studies 2 and 3, different-voice, different-box, praise condition)

Study	Tutor Configuration	Tester Configuration	Evaluator Configuration	Valence of Evaluation	Questionnaire Given By
1. Politeness	C1 / V1	C1 / V1	C1 / V1	Praise only	C1 / V1 or C2 / V2 or Paper & pencil
2. Voice and Box	C1 / V1	C2 / V2	C1 / V1 or C3 / V3	Praise or Criticism	Paper & pencil
3. Voice or Box	C1 / V1	C2 / V2	C1 / V1 or C1 / V3 or C3 / V1 or C3 / V3	Praise or Criticism	Paper & pencil
4. Gender	C1 Voice M or F	C2 No voice	C3 Voice M or F	Mixed Praise & Criticism	Paper & pencil
5. Computer / Programmer	C1 / V1	C2 / V2	C1 / V1 and Computer or Programmer or Computer / "I"	Praise and Criticism	Paper & pencil

C = computer; V = voice

Table 1: Summary of Experimental Conditions for All Studies



The **practice session** provided a brief introduction to the use of the interface controls and to the voice output. The **tutoring session** provided subjects with a total of 25 to 30 facts, which ostensibly were selected from a large database and were based on the subjects' knowledge of the topic, such as mass media, computers, social customs, and love and relationships. The **testing session** involved a 15-item, five-alternative multiple choice test. Although subjects were told in the instructions that a total of 15 questions would be randomly chosen from a set of 2500 possible questions, in fact, all subjects were given the same 15 questions. During the **evaluation session**, a computer evaluated the performance of the computerized tutoring session. (See Figure 1). After the evaluation, subjects answered a **questionnaire** to make assessments of the tutoring, testing, and evaluation sessions. All of the questionnaire items were measured on ten-point Likert scales.

### STUDY 1: COMPUTERS AND POLITENESS

Will people apply politeness norms to computers? This experiment demonstrates that people apply the following social rule to computers: "Direct, as opposed to indirect, requests for evaluations elicit more positive responses and more homogeneous responses." In other words, users asked by a computer about its own performance will feel compelled to be more positive than will users asked about the computer by an independent source.

#### Study 1 Method

In this experiment, 33 subjects used one or two computers. All subjects used a single computer for the tutoring, testing, and evaluation sessions. The experimental manipulation was the locus of assessment. The questionnaire was administered in one of three ways: on the same computer as the tutoring, testing, and evaluating sessions; on a paper and pencil questionnaire (P&P); or on a second, distinct computer (see Table 1). We expected that in the same-computer condition, subjects would treat the computer as if it were asking for a direct evaluation of itself and would therefore be positive and highly constrained in their assessments. However, the P&P questionnaire should be treated by subjects as an independent request for an evaluation; hence, subjects should therefore be more honest and varied in their responses. If the distinct computer operates like the P&P condition, then we can rule out medium effects. The dependent variables in this study were the subject's perceptions of the computer and the variance in subjects' perceptions.

#### Study 1 Results

Although subjects indicated in debriefing that norms of politeness do not apply to interactions with computers, they did apply them in evaluating the computer. As predicted, in assessing the computer's performance in the same-computer condition, subjects said that the tutoring was more friendly (based on a five-item index;  $p < .01$ ) and more competent (based on a four-item index;  $p < .02$ ) than in the P&P assessment. Similarly, using the same items for the indices, the computer-based evaluation session was said to be more friendly ( $p < .001$ ) and more competent ( $p < .01$ ) when the computer asked about itself. Also

consistent with our hypotheses, subjects' responses were significantly less varied in the same-computer condition ( $p < .001$ ), again providing evidence that the direct evaluation rule was employed.

Using a distinct computer for the subjects' assessment, rather than a P&P questionnaire, did not significantly alter subjects' responses. In short, there were the same significant differences between the same-computer and different-computer assessment sessions, while there were no significant differences between the different computer and the P&P conditions. This confirms the social explanation and contradicts the modality explanation.

### STUDY 2: "SELF" VS. "OTHER:" BOX AND VOICE

Will users apply the notions of "self" and "other" to computers? This experiment demonstrates that people apply the following social rules to computers: "Other-praise is more valid than self-praise," "Other-praise is friendlier than self-praise," "Self-criticism is friendlier than other-criticism," "Other-critics are more intelligent than other-praisers," and "A praised performance is superior to a criticized performance."

#### Study 2 Method

In this experiment, 44 subjects used two or three computers in a 2 (praise or derogation) by 2 (same voice and box / different voice and box), between-subjects design. All experimental manipulations were introduced during the evaluation session. In the "same voice and box" condition, the evaluation session was on the same computer and in the same voice as the tutoring session (only the testing session was on a distinct computer). In the "different voice and box" condition, the evaluation session was on a different computer and in a different voice than the tutoring session (see Table 1). In the "praise" condition, the computer described the performance of the tutoring session favorably for 12 of the 15 questions; for "criticism," the performance was described negatively for 12 of the 15 questions. The questionnaire for assessing the tutoring, testing, and evaluation sessions were all paper and pencil. The dependent variables were the subject's perception of the tutor, the test, the evaluation, and his or her own performance.

#### Study 2 Results

As in the politeness study, subjects used inappropriate social rules in assessing machine behavior. Subjects responded consistently with a belief that a different computer with a different voice is a distinct social actor. When assessing the praise condition with different vs. same box, they applied the social rules that praise of others is more accurate and friendly than praise of self. Specifically, when the praised evaluation was on a different box and in a different voice, subjects said the tutor performed better (12 of 14 items;  $p < .01$ ), the praise was more accurate ( $p < .05$ ), and the praise was more friendly ( $p < .01$ ).

Also consistent with treating a different box with a different voice as a distinct social actor, subjects used the rule that other-critics are less friendly but more intelligent than other-praisers (see Figure 2): In the derogation condition,



subjects assessed the different box and voice evaluation as less friendly ( $p < .001$ ) and more intelligent (see Figure 3;  $p < .03$ ).

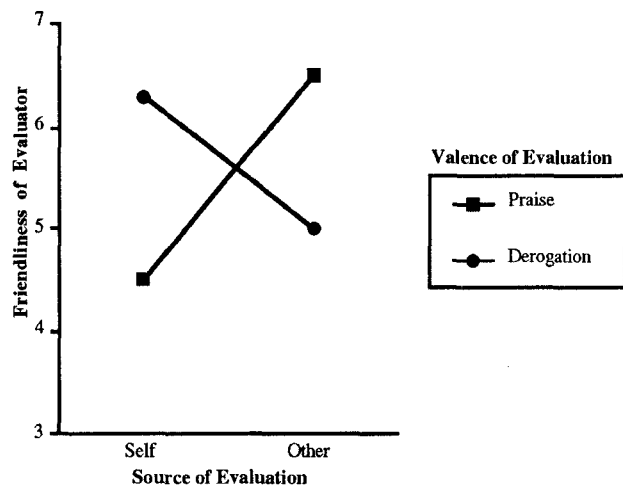


Figure 2: Perceived Friendliness as a Function of the Source and Valence of the Evaluation ( $N = 44$ )

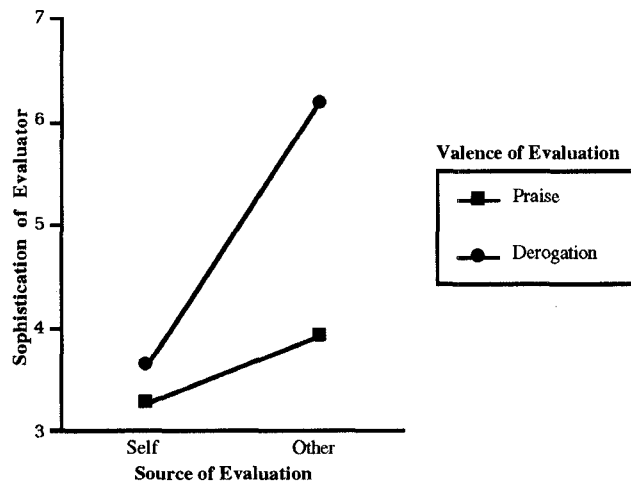


Figure 3: Perceived Sophistication as a Function of the Source and Valence of the Evaluation ( $N = 44$ )

Finally, the rule that a praised performance is better than a criticized performance was applied by subjects, as a praised tutor was evaluated as having performed better (14 of 14;  $p < .001$ ) than a criticized tutor. Subjects even perceived that they answered more questions correctly in this condition ( $p < .001$ ), though in fact they did not. Thus, notions of “self” and “other” are applied to computers.

**STUDY 3: “SELF” VS. “OTHER:” BOX OR VOICE**

On what basis do users distinguish computers as “self” or “other” — the voice or the box? The social rules tested by this experiment are the same as those in Study 2; however, this study takes the question one step further by pinpointing the locus of the self / other attribution.

**Study 3 Method**

This experiment, which involved 66 subjects, is an extension of the previous study. It is a 2 (praise / criticism) by 2 (same voice / different voice) by 2 (same box /

different box), between-subjects design. The experiment is identical to Study 2, except that the voice in which the evaluation is presented as well as the box on which the evaluation occurs vary independently. That is, eight distinct conditions are possible during evaluation (see Table 1). As in Study 2, the dependent variables were the subject’s perception of the tutor, the test, the evaluation, and his or her own performance.

**Study 3 Results**

Subjects responded to different voices as if they were distinct social actors, and to the same voice as if it were the same social actor, regardless of whether the different voice was on the same or different computer. That is, different voices on the same computer elicit the same responses as different voices on a different computer. Specifically, in both the same and different computer conditions, different-voice subjects perceived the evaluation session to be significantly more accurate ( $M = 5.3$ ) than did same-voice subjects ( $M = 4.2$ ,  $F(80,1) = 4.66$ ,  $p < .03$ ). This result is consistent with subjects treating different voices as distinct social actors and treating the same voice as the same social actor. Different-voice subjects also perceived the tutoring session to be fairer ( $M = 6.4$ ) than did same-voice subjects ( $M = 5.5$ ,  $F(80,1) = 4.07$ ,  $p < .05$ ). Finally, when asked to assess the tutor’s performance relative to the evaluation, different-voice subjects perceived less absolute difference ( $M = 1.3$ ) than did same-voice subjects ( $M = 1.9$ ,  $F(80,1) = 7.6$ ,  $p < .01$ ).

In assessing the friendliness of the computers, subjects applied the social rules, “Praise of others is friendlier than praise of self,” and “Criticism of self is friendlier than criticism of others” to the computers. Figure 4 presents the mean perceived friendliness for each of the eight cells:

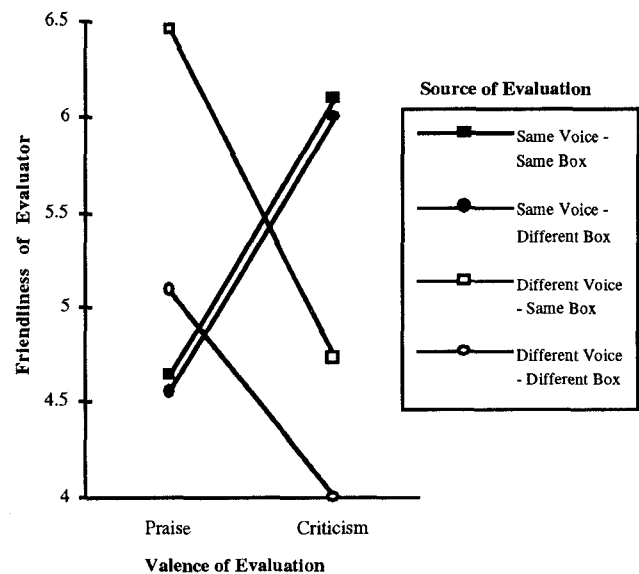


Figure 4: Perceived Friendliness as a Function of Source and Valence of Evaluation

There was a significant interaction between the valence of the evaluation and whether the tutoring and evaluation



session were in the same voice or different voices ( $F(1,80) = 10.31, p < .002$ ), based on a full factorial analysis of variance. For praise, same-voice subjects found the evaluation to be significantly *less* friendly than did other-voice subjects ( $t(42) = 1.9, p < .03$ ); for criticism, same-voice (modesty) subjects found the evaluation to be significantly *more* friendly than did other-voice subjects ( $t(42) = 2.7, p < .005$ ). There were no significant differences for same vs. different box for either accuracy or friendliness.

Thus, voice, and not box, is the primary determinant of the locus of social attributions toward computers.

#### STUDY 4: ENGENDERING COMPUTERS

Will users apply gender stereotypes to computers? [9] This experiment demonstrates that people apply the following social rules to computers: "Praise from males is more convincing than praise from females," "Males who praise are more likable than females who praise," and "Females know more about love and relationships than males."

##### Study 4 Method

In this experiment, 48 subjects, 24 males and 24 females, used three computers. In order to study the effect of the gender of voices, the testing session had no voice and the tutoring and evaluation sessions were given either a female or male voice. These factors were varied independently, and were counterbalanced for subject gender; thus, six male and six female subjects participated in each of four possible combinations of voice gender (see Table 1). The topics for the tutoring and testing were love and relationships, mass media, and computers. The dependent variables were the subject's perceptions of the tutor and evaluator.

##### Study 4 Results

As in the previous studies, subjects applied the social rules in question, even though they knew the rules to be inappropriate. Using the rule that praise from males is more convincing than praise from females, tutors (whether male- or female-voiced) who were evaluated by a male voice were seen as more assertive ( $p < .01$ ), more dominant ( $p < .1$ ), more forceful ( $p < .05$ ), more affectionate ( $p < .05$ ), more cheerful ( $p < .1$ ), more sympathetic ( $p < .05$ ), and warmer ( $p < .1$ ) than tutors evaluated by female voices.

Similarly, demonstrating that subjects use the social rule, "Males who praise are more likable than females who praise," evaluators with a male voice were seen as more dominant ( $p < .1$ ), more assertive ( $p < .1$ ), more forceful, ( $p < .05$ ) more sympathetic ( $p < .05$ ), and warmer ( $p < .05$ ) than evaluators with a female voice.

Also, in line with the belief that females know more about love and relationships than males, female tutors of love and relationships were evaluated as more sophisticated ( $p < .1$ ) and having chosen better ( $p < .1$ ), broader ( $p < .1$ ), and less-known facts ( $p < .05$ ) than male tutors.

Another interesting set of interactions emerged in the conditions in which the voices of the tutor and evaluator

were of different genders. In both the male tutor / female evaluator and the female tutor / male evaluator conditions, both the style of tutoring and the style of evaluation was seen as more similar to the subject's own style (both  $p < .1$ ). In addition, in these different-gender conditions, facts about love and relationships were seen as more informative ( $p < .05$ ), better-chosen ( $p < .05$ ), and broader ( $p < .05$ ) than in the same-gender conditions.

Thus, gender stereotypes applied in human-human interaction are applied in human-computer interaction with computers employing gendered voices.

#### STUDY 5

When people respond socially to computers, is it because they feel that they are interacting with the computer or with some other agent? [10] This study demonstrates that users do make social attributions toward the computer itself, and provides some evidence that the computer referring to itself as "I" is not necessary to generate such attributions.

##### Study 5 Method

In this study, 33 subjects used two computers. The protocol differed from previous studies in that subjects experienced the tutoring / testing / evaluating cycle twice, rather than only once. In the first cycle, the evaluations were generally positive; in the second, they were generally negative. Tutoring and evaluation took place on one computer; the test was administered by a distinct computer (see Table 1).

The experimental manipulation was the way in which the tutor / evaluator computer was referred to by the computer itself and by the experimenter. In the "computer" condition, both experimenter and computer referred to the computer as "this computer" or "the computer." In the "computer - 'I'" condition, the computer referred to itself as "I," but the experimenter referred to it as "the computer." In the "programmer" condition, the computer referred to itself as "I," but the experimenter referred to it as "the programmer." The dependent variables were subject's perceptions of the computers, as measured by 170 different questionnaire items.

##### Study 5 Results

Subjects in the computer condition found the computer to be generally more capable, more likable, and easier to use than the computer in the programmer condition. Of 170 comparisons of mean responses between these groups, 167 of the differences were in the direction of greater performance and greater liking, and 60 of these differences were significant. The computer in the computer - 'I' condition was also better liked and considered more capable by subjects than was the computer in the programmer condition. Of 170 mean comparisons, 148 were in the appropriate direction, and 40 of these differences were significant. Subjects did not perceive the computer and computer - 'I' conditions to be substantially different, rating the computer condition higher on only 95 of 170 items, and none of these differences were significant. No differences were found between the positive and negative evaluation conditions.



Thus, computer users who respond socially to computers do not feel as though they are interacting with a programmer, but rather attribute socialness directly to the computer itself. Computer self-reference using “I” is not essential to generating this response.

### **THEORETICAL IMPLICATIONS**

This section summarizes the theoretical implications of the experiments described above.

#### **Study 1**

- Social norms are applied to computers.

#### **Study 2**

- Notions of “self” and “other” are applied to computers.

#### **Study 3**

- Voices are social actors.
- Notions of “self” and “other” are applied to voices.

#### **Study 4**

- Computers are gendered social actors.
- Gender is an extremely powerful cue.

#### **Study 5**

- Computer users respond socially to the computer itself.
- Computer users do not see the computer as a medium for social interaction with the programmer.

#### **All Studies**

- Primitive cues are powerful.
- Social responses are automatic and unconscious.
- Findings in social psychology are relevant to human responses to computers
- Human-computer interaction is social-psychological.
- Experimental paradigms are interchangeable.

### **DESIGN IMPLICATIONS**

This section summarizes the design implications of the experiments described above.

#### **Study 1**

- Usability testing should not be same-machine based.

#### **Study 2**

- Modesty is a complex phenomenon.
- Integration is highly consequential.
- Uniformity of interface is double-edged.
- Social cues need not be heavy-handed.

#### **Study 3**

- Choice of voices is highly consequential.

- Agent integration/differentiation is easily created.
- Agent integration/differentiation is powerful.
- Social actors can be easily portrayed on multiple computers. Cross-computer continuity of social actors is easily obtained.
- Uniformity of interface is double-edged.

#### **Study 4**

- Gender of voices is highly consequential.

#### **Study 5**

- Computers need not refer to themselves as “I” to generate social responses.

### **CONCLUSIONS**

We have presented five studies which show that experienced computer users do in fact apply social rules to their interaction with computers, even though they report that such attributions are inappropriate. These social responses are not a function of deficiency, or of sociological or psychological dysfunction, but rather are natural responses to social situations. Furthermore, these social responses are easy to generate, commonplace, and incurable.

We have shown that the human-computer relationship is fundamentally social. These results suggest that many other principles drawn from the extant literature in social psychology, communication, and sociology are relevant to the study of human-computer interaction and have clear implications for user interface design.

Traditionally, when interface agents have been created, they have been endowed with faces, personalities, and a rich human representation [11,12]. The present research suggests that low-overhead agents can be easily produced and can generate a wide range of social responses., and suggests that concern with the inability to create a photo-realistic, full-motion video, or other high-bandwidth representation may be highly overrated.



### ACKNOWLEDGMENTS

This research was funded by US West Advanced Technologies. Additional funding was provided by Apple Computer, IBM, Microsoft, and US West, and through a National Science Foundation Graduate Fellowship.

The authors would like to thank Byron Reeves for his general contributions to this research. In addition, the following students participated in the design and implementation of these studies: [Study 1] Mercy Malick, Scott Reiss, and Douglas Wade; [Study 2] D. Christopher Dryer and Lisa Henriksen; [Study 3] Heidi Reeder; [Study 4] Nancy Green and Mercy Malick; and [Study 5] Shyam Sethuraman, and Albert Cho.

Thanks to Lisa Seaman for her artistic efforts on Figure 1.

### REFERENCES

1. Turkle, S. *The Second Self: Computers and the Human Spirit*. Simon & Schuster, New York, NY, 1984.
2. Rafaeli, S. *Interacting with Media: Para-social Interaction and Real Interaction*. Unpublished doctoral dissertation, Stanford University, 1985.
3. Cooley, C. H. *Social process*. Southern Illinois University Press, Carbondale, IL, 1966.
4. Nass, C., and Steuer, J. Anthropomorphism, Agency, & Ethopœia: Computers as Social Actors. *Human Communication Research*, 19(4) (June 1993), 504-527.
5. Steuer, J. *It's Only Talk: Speech as A Possible Determinant of the Social Categorization of Computers*. Unpublished manuscript, Stanford University, 1990.
6. Nass, C., Steuer, J., Tauber, E., and Reeder, H., Anthropomorphism, Agency, & Ethopœia: Computers as Social Actors. Presented at INTERCHI '93; Conference of the ACM / SIGCHI and the IFIP; Amsterdam, the Netherlands — April, 1993.
7. Nass, C., Steuer, J., Henriksen, L., and Dryer, D. C. Machines, Social Attributions, and Ethopœia: Performance Assessments of Computers Subsequent to "Self-" or "Other-" Evaluations. *International Journal of Man-Machine Studies*, In Press.
8. Winograd, T. and Flores, F. *Understanding Computers and Cognition*. Addison-Wesley, Reading, MA, 1987.
9. Green, N. Can Computers Have Genders? Paper presented to the Communication and Technology Division, Conference of the International Communication Association; Washington, DC — May, 1993.
10. Sethuraman, S. S. Communicator, Programmer, or Independent Actor: What Are Computers? Paper presented to the Communication and Technology Division, Conference of the International Communication Association; Washington, DC — May, 1993.
11. Laurel, B. Interface Agents: Metaphors with Character. In *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, MA, 1990.
12. Oren, T., Salomon, G., Kreitman, K., & Don, A. Guides: Characterizing the Interface. In *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, MA, 1990.