UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Anthropomorphic Agents as a User Interface Paradigm: Experimental Findings and a Framework for Research

Permalink

Journal

ISSN

Authors

Catrambone, Richard
Stasko, John
Xiao, Jun

Publication Date

2002

Peer reviewed

Anthropomorphic Agents as a User Interface Paradigm: Experimental Findings and a Framework for Research

Richard Catrambone (rc7@prism.gatech.edu)
School of Psychology, Georgia Institute of Technology, 274 5th Street
Atlanta, GA 30332-0170 USA

John Stasko (stasko@cc.gatech.edu)
Jun Xiao (junxiao@cc.gatech.edu)
College of Computing, Georgia Institute of Technology
Atlanta, GA 30332-0280 USA

Abstract

Research on anthropomorphic agent interfaces has produced widely divergent results. We suggest that this is due to insufficient consideration of key factors that influence the perception and effectiveness of agent-based interfaces. We propose a framework for studying anthropomorphic agents that can systematize the research. The framework emphasizes features of the agent, the user, and the task the user is performing. Our initial experiment within this framework manipulated the agent's appearance (lifelike versus iconic) and the nature of the user's task (carrying out procedures versus providing opinions). We found that the perception of the agent was strongly influenced by the task while features of the agent that we manipulated had little effect.

Introduction

If you could ask for assistance from a smart, spoken natural language help system, would that be an improvement over an on-line reference manual? Presumably the answer, in most cases, is yes, for two reasons. First, the spoken natural language aspect would allow you to speak your questions rather than having to type them. Generally this is a faster approach for most people. Second, the smart aspect would improve the chance of the help system finding the information you want even if you do not state the query using the correct or most appropriate terms.

The state of the art in this style of interface is a human user consultant. Does it matter that the user consultant has a face and that the face can have expressions and convey a personality? Would a face affect you in terms of your comfort and satisfaction with the interaction? Would the presence of a face make the help or advice you receive more persuasive? The answers to such questions have implications for the design of training systems, customer service, information kiosks, and many other applications.

Many people believe that anthropomorphic computer interfaces have great potential to be beneficial for a number of reasons. Agents could act as smart assistants, much like travel agents or investment advisors, helping people manage the ever-growing amount of information encountered today (Lyman & Varian, 2002). Further, a conversational interface appears to be a natural dialog style in which the user does not have to learn complex command structure and functionality (Laurel, 1990).

An anthropomorphic interface could use intonation, gaze patterns, and facial expressions, in addition to words, for conveying information and affect. The human face seems to occupy a privileged position for conveying a great deal of information, including relatively subtle information, efficiently (Fridlund & Gilbert, 1985). Anthropomorphic interfaces could make a computer more human-like, engaging, entertaining, approachable, and understandable to the user, thus harboring potential to build trust and establish relationships with users, and make them feel more comfortable with computers.

These potential advantages are balanced by strong negatives. Anthropomorphic agent interfaces are viewed by some researchers as being impractical and inappropriate. Current speech recognition, natural language understanding, and learning capabilities of computers still fall far short of any human assistant.

More specifically, it has been proposed that agent systems disempower users by clouding issues such as who is responsible for a system's actions (Lanier, 1995). Others feel that user interfaces are more beneficial when they clearly reflect the commands available to a user and present the objects that a user can act upon (Shneiderman, 1997). Furthermore, critics argue that agent interfaces may mislead both users and designers, increase user anxiety, reduce user control, undermine user responsibility, and destroy a user's sense of accomplishment (Shneiderman & Maes, 1997). Many current anthropomorphic or personified interfaces are viewed as being annoying, silly characters who hinder rather than enhance productivity (e.g., the Microsoft Paper Clip).

Although strong opinions have been voiced on both sides of this issue, relatively little careful empirical research on anthropomorphic interfaces has been conducted, and the results from this research have been contradictory or equivocal (Cassell, 2000).

Our goal is to develop a framework to systematically evaluate and understand the autonomous agent as a user interface paradigm. The present paper outlines the framework and an initial study that examines two issues within this framework. The first is whether the degree to which an interface agent is anthropomorphic has a measurable effect on users. Note that anthropomorphism is not a dichotomy but rather a continuum. One can think of interfaces with full fidelity video or 3D images of people to

more caricature-style characters to 2D cartoons of people or personified characters such as dogs or toasters.

The second issue is to what extent the nature of the task will influence a user's perception of an agent. Some tasks might be more likely to induce a user to imbue the agent with human-like qualities (such as if the user had to engage the agent in a debate) while other tasks might lead the user to view the agent simply as a reference tool (e.g., for providing reminders of keystrokes for a software application) with no "individuality."

Related Work

A few studies have revealed that anthropomorphic agents are attention-grabbing and people make natural assumptions about the intelligence and abilities of those agents. King and Ohya (1996) found that a dynamic 3D human form whose eyes blinked was rated more intelligent than any other form, including non-blinking 3D forms, caricatures, and geometric shapes.

One common trend discovered in studies is that anthropomorphic interfaces appear to command people's attention, both in positive and negative senses. Takeuchi and Nagao (1995) created conversational style interaction systems that allowed corresponding facial displays to be included or omitted. According to their metrics, the conversations with a face present were more "successful." Across two experiments they found that the presence of a face provided important extra conversational cues, but that this also required more effort from the human interacting with the system and sometimes served as a distraction.

Other studies have shown that the attention garnered by an anthropomorphic interface had a more positive effect. Walker, Sproull, and Subramani (1994) found that people who interacted with a talking face spent more time on an on-line questionnaire, made fewer mistakes, and wrote more comments than those who answered a text questionnaire. Koda (1996) created a Web-based poker game in which a human user could compete with other personified computer characters including a realistic image, cartoon male and female characters, a smiley face, no face, and a dog. She gathered data on people's subjective impressions of the characters and found that people's impressions of a character were different in a task context than in isolation and were strongly influenced by perceived agent competence.

An influential body of related work is that of Nass and his colleagues. Their efforts focus on the study of "Computers as Social Actors." They have conducted a number of experiments that examined how people react to computer systems and applications that have certain personified characteristics (Nass, Isbister, & Lee, 2000; Nass, Steuer, & Tauber, 1994; Rickenberg & Reeves, 2000). Their chief finding is that people interact with and characterize computer systems in a social manner, much as they do with other people. Furthermore, they suggest that findings in the social psychology literature (e.g., individuals with similar personalities tend to get along

better than do those with different personalities) apply even when one of the two participants is a machine.

The studies cited above, and others, suggest that people are inclined to attribute human-like characteristics to agents and that a variety of factors might influence how positively the agents are viewed. Dehn and van Mulken (2000) provide a more extensive review of this literature.

A Framework for Research on Anthropomorphic Interface Agents

To effectively and systematically investigate the use of anthropomorphic interface agents, one needs to consider the key factors that will affect the usefulness of such interfaces. We propose an investigative framework composed of three key components: characteristics of the user, attributes of the agent, and the task being performed.

We believe that serious empirical study in this area must systematically address each of these factors and understand how it affects human users. Below, we provide examples of individual variables within each factor that could potentially influence user performance and impressions.

Factor 1: Features of the User

Potential users vary, of course, in many ways. However, there are certain features that may be quite likely to affect how useful a user finds an agent. These features include:

Personality: Researchers have identified what are referred to as the "Big Five" traits that seem to be quite useful in describing human personalities: extraversion, openness, agreeableness, neuroticism, and conscientiousness (e.g., McCrae & Costa, 1987). While any such breakdown is debatable, it seems reasonable to examine whether users' positions on these, or other, trait dimensions predicts how they will respond to agents.

Background Knowledge: A user who has a good deal of background knowledge in a domain might prefer an agent that is reactive and that the user can call upon when he or she needs some low-level bit of information or has a low-level task that needs to be done. Conversely, a user who is learning how to carry out tasks in a particular domain might welcome strategy advice from an agent, particularly if the agent can analyze the strategy and provide reasons for why the strategy might be altered.

Other Variables: Other user-related variables include gender, age, and computer experience.

Factor 2: Features of the Agent

Fidelity: Earlier studies suggest that more realistic-appearing, 3D human representations are perceived as being more intelligent, which could be viewed positively or negatively. However, realistic-appearing agents are more difficult to implement, so if user performance is improved by the presence of an agent, but does not vary according to appearance, simpler caricature style characters would be advantageous.

Presence: Is an agent's face always present on the screen or does the agent only appear when it is engaged in a dialog

with the user? One might hypothesize that an ever-present agent would make users uneasy by producing an effect of being watched or evaluated all the time.

Role: Should an agent act as a partner in the task or should it contribute only in clearly specified ways? For instance, an agent might be able to offer strategy guidance for design tasks. Alternatively, it might provide only lower-level procedural "how to" information.

Initiative: Should an agent proactively make suggestions and offer guidance or should it respond only when directly addressed? A proactive agent might be viewed as being "pushy" and might bother users, or it could be viewed as being extremely helpful and intelligent if it acts in situations in which the user is unsure of how to proceed or is so confused that he or she is unable to form a coherent help request.

Other Variables: Other agent-related variables to consider are expressiveness, speech quality, "gender," "personality," and competence.

Factor 3: Features of the Task

Tasks can vary in a variety of ways. Some tasks can be opinion-like (e.g., choosing what to bring on a trip) while others are more objective (e.g., solving a puzzle) in terms of assessing the quality of a solution. Some involve a good deal of high-level planning (e.g., writing a talk) while others are more rote (e.g., changing boldface words into italics). Tasks might be classified along some or all of the dimensions listed below:

Objectiveness: The situation might be an opinion-based one in which the user is seeking advice and recommendations on some topic (e.g., which items to pack for a trip to Europe). Alternatively, the user might be carrying out an objective task such as acquiring facts (e.g., finding the keystroke combination for a particular command in a software application).

Intent: The user could have a learning goal or alternatively may be carrying out a set of steps in a familiar domain. In the latter, the user might want help with low-level details whereas in the former the user is looking for guidance as to the structure of the domain.

Other Variables: Other task-related variables to consider are domain, degree of time pressure, duration, and consequences of the quality of task performance.

The number of variables within each factor is certainly larger than the number we have identified here. No doubt these factors will also interact. For instance, a novice attempting to carry out a task in a particular domain might welcome proactive comments/advice from an agent while someone with more experience could get annoyed.

With respect to measuring the usefulness of an agent, we have to consider which dependent measures are most appropriate. Towards the more objective end, a user's performance on a task in terms of accuracy and time--when such measures are meaningful--can give one indication of usefulness. Thus, time and errors would be appropriate measures for a text-editing task. Towards the more

subjective end, a user is likely to have a number of affective reactions to an agent. These reactions might manifest themselves in terms of how much users liked the agent, how intrusive they found the agent, how they perceived the agent's personality, and how willing they are to use the agent in the future. We can certainly assess a user's liking and satisfaction towards an agent, but if the user can carry out the tasks more effectively with the agent regardless of liking and satisfaction, then how important are those variables? On the other hand, long-term use of an agent might be predicted by liking and satisfaction.

The likelihood of a user following an agent's advice might be another interesting measure of the usefulness of an agent. While advice-following would certainly be at least partly a function of the quality of the advice, it will also be impacted by how the user feels about the agent (how many children ignore the advice of their parents merely because it is the parents giving the advice?).

Experiment

Overview

One fundamental issue in the quality of agent interfaces is competence (Maes, 1994). It appears obvious that perceptions of anthropomorphic agent interfaces will be strongly influenced by the competence of the supporting software system and the quality of the replies and suggestions made by the agent. We chose to factor out competence as an influence. If our experiments uncover that people's performance is not enhanced and they dislike anthropomorphic user interfaces even though the system is competent, then that is an important and strong result that other researchers and developers need to understand. To remove competence as a factor, we employed a "Wizard of Oz" (Dahlback, Jonsson, & Ahrenberg, 1993) methodology (described below).

The experiment manipulated the agent fidelity and the task objectiveness variables because prior work and our framework suggest they seemed likely candidates to have an affect on the perception of agents. Usefulness was evaluated via both the performance and satisfaction dimensions. We hypothesized that user reactions to the agent would vary as a function of the objectiveness of task. A task that required the user to debate the merits of his or her opinion (about items to pack on a trip) might lead the user to feel the agent had more of a personality (for good or for bad) compared to a task in which the user made use of the agent more as a reference tool (i.e., reminding the user of keystroke commands for a text editor). We also hypothesized that users might find the agent to be more useful in its role as a reference source rather than as an entity that provides opinions. Finally, we expected that the more life-like the agent appeared, the more likely the user might be to ascribe qualities such as personality and intelligence to the agent, but objective performance would likely not be affected by appearance.

Participants

Thirty-nine undergraduates participated for course credit and were randomly assigned to conditions. Participants had a variety of majors and computer backgrounds.

Procedure and Design

Participants were run individually using a computer equipped with a microphone and speaker. Participants performed two tasks: a travel task and an editing task. The travel task was chosen to be a type of creative, opinion-based task in which interacting with an agent might be viewed as an opportunity to think more deeply about the task by discussing points of view about the importance of travel items. The editing task was chosen to represent an opportunity to use an agent primarily as a reference source rather than as a guide or teacher.

The travel task involved a hypothetical situation in which the participant had a friend who was flying overseas on his first international trip. The task was to recommend six items for the person to take with him from a pool of 12 items and to rank the six items in order of importance.

After the participant did the initial ranking using a simple software interface, a computer agent who supposedly had knowledge about international trips appeared. The agent made a predefined set of suggestions in which it recommended changing the rankings of four of the six choices and it agreed with the ranking of two other items. For example, the agent first suggested promoting the person's fourth item to the first position, demoting the first item but keeping it in the top six. The agent explained the reasoning for its suggestion at every stage and asked the participant what he or she thought about the suggestion. After the participant responded to the agent's comment on a particular item, the agent would say one of several conversational conventions (e.g., "OK, let's continue") so that it could move on to the next suggestion. After the agent finished providing feedback on the rankings, the original rankings were displayed on the screen and the participant was given the opportunity to change the rankings. After doing the re-ranking, participants filled out a questionnaire about the agent and were asked a few questions about the agent and related issues by the experimenter.

The editing task required participants to use an unfamiliar text editor to modify an existing document by making a set of prescribed changes to the document. Participants first viewed a short video that described the various functions (e.g., copy, paste) and the specific key combinations needed to issue the commands. Participants were then shown a marked-up document that required a set of changes such as deletions, insertions, and moves, and they were instructed that if at any time they could not remember the keystrokes for a particular function, they could ask the agent for help. Pilot testing was conducted to ensure that the number of commands was sufficiently large so that participants would be likely to need to ask the agent for help. After completing the editing tasks, participants again filled out a questionnaire about the agent and answered questions from the experimenter.

The agent was controlled through a Wizard of Oz technique. One experimenter was in the room with the participant to introduce the experimental materials, and a second experimenter was in an adjacent room, monitoring the questions and responses made by the participant. The second experimenter insured that the agent responded in a consistent manner using a prepared set of replies.

Two between-subjects variables were manipulated: type of agent (animated, stiff, iconic) and task order (travel task then editing task or vice versa). The left side of Figure 1 shows the face of the agent in the animated and stiff conditions. The animated agent (donated by Haptek Corp.) was 3D, with a female appearance--though somewhat androgynous--that blinked, moved its head, and produced certain facial expressions in addition to moving its mouth in synchronization with the synthesized voice. The stiff agent had the same face as the animated agent but moved only its mouth. The iconic agent (see the right side of Figure 1) was a light-bulb icon that had arrows appear whenever it spoke.

One design issue about this experiment should be flagged. Although our key task manipulation was the "objectiveness" of the task (i.e., the travel task being less objective and the editing task being more objective), the nature of the agent also was varied as a function of the task. The agent was completely reactive in the editing task; it provided information only when requested. However, in the travel task the agent provided feedback regardless of the participants' desire. A cleaner version of the experiment would have been to hold the "nature" of the agent constant across the tasks. We allowed this confounding to occur here because were interested in getting participants' reactions to certain human-like attributes of the agent but did not have the resources to run the additional conditions that would have been required to completely cross this factor with the task and appearance manipulations. In future work we plan to systematically investigate this reactive/proactive dimension.
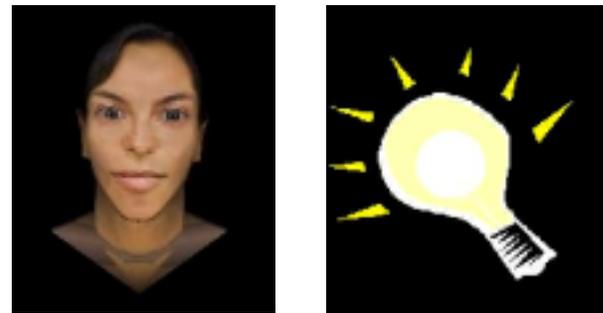
Figure 1: Appearance of Agent in Animated and Stiff Conditions (left) and Iconic Condition (right).

Measures

Both objective and subjective measures were used. One objective measure was, for the travel task, whether

participants changed their rankings as a function of the agent's feedback. For the editing task we measured how long it took participants to complete the edits. The primary subjective variables were the responses to the individual items in the questionnaires and the answers to the questions posed by the experimenter. The questionnaire items used a five-point Likert scale (1 = strongly agree, 5 = strongly disagree) that addressed a number of qualities of the agent (see Table 2). The questions posed by the experimenter were open-ended and provided participants an opportunity to give their impressions about the agent's personality, helpfulness, and intelligence.

Results

In the data analyses we found that the task order manipulation did not have an effect, so in the interest of simplicity we will collapse across that factor in the presentation and discussion of the results.

Performance Measures. With respect to more objective measures, Table 1 shows that participants were more likely to change the rankings of items that the agent disagreed with compared to items that the agent agreed with, $F(1, 36) = 38.48$, $MSE = .07$, $p < .0001$). There was no effect of type of agent, $F(2, 36) = 0.9$, $MSE = .11$, $p = .42$. There was no interaction, $F(2, 36) = 1.25$, $p = .30$.

Table 1: Proportion of Travel Items with Changed Rankings as a Function of Type of Agent & Agent Advice.

	Animated ($n =14$)	Stiff ($n =12$)	Iconic ($n =13$)
Suggested Change	.82	.90	.77
Keep Rank	.57	.42	.38

The time (in seconds) to do the editing task did not differ as a function of agent (animated: 714.8, stiff: 568.7, iconic: 671.1); $F(2, 31) = 1.78$, $MSE = 37637.22$, $p = .19$ (5 participants did not do the editing task).

Questionnaire Responses. Table 2 shows the mean responses to the questionnaire items for the different agent conditions after the travel and editing tasks (there were 5 participants who did not do both tasks and they are excluded from Table 2). There was no effect of agent type for any of the questions. For two of the items, worthwhile and intrusive, there was an effect of task (worthwhile: $F(1, 31) = 15.68$, $MSE = .45$, $p = .0004$; intrusive: $F(1, 31) = 20.28$, $MSE = .23$, $p = .0001$). The agent was rated more worthwhile and less intrusive after the editing task compared to the travel task. These results make sense. First, the editing task required most participants to rely heavily on the agent to remind them of commands, thus making the agent seem worthwhile. Second, the uninvited critique of participants' rankings of travel items could certainly have seemed intrusive.

While group differences did not exist on most of the questionnaire items, it is interesting that for most items, the average response tended to be in the positive direction. Participants felt positively, on average, about the agent.

Interview Responses. While participants made a number of interesting and insightful comments about the agent in response to questions from the experimenter, a simple tally of responses shows reactions to the agent that again varied as a function of task. Virtually all participants found the agent helpful for both tasks. Participants were much less likely to consider the agent to have a personality after doing the editing task compared to the travel task. This makes sense because the agent was merely providing subjects with information on commands in the editing task. In the travel task the agent expressed its "opinions."

Finally, it is worth noting that in general the agent was perceived as more intelligent after the travel task than after the editing task. At one level this seems odd because the agent had all the answers for the editing task. However, as demonstrated by some participants' comments, the agent was perceived as very limited in the editing task; it knew about editing commands and probably little else (despite the fact that it also appeared to understand spoken language!). In the travel task though it presumably gave the impression of having sufficiently deep knowledge about travel such that it could give feedback on the importance of various items one might take on a trip. While some of the participants' responses to the agent indicated that they disagreed with its suggestions, they appeared to believe that the suggestions were at least thoughtful.

Discussion

In addition to the results reported above, we learned a great deal by observing participants' behaviors and responses in the sessions. One key question we had was how would the participants interact with the agent in the two different tasks. In the editing task, participants seemed very comfortable asking the agent for assistance. Participants requested help an average of 6.5 times. However, in the travel task participants seemed reluctant to engage the agent in a dialog. Only a few replied with more than a few words when the agent engaged them. There was clearly awkwardness to the interaction.

The agent's social abilities and personality (or lack thereof) were noted by a number of the participants. In the travel task, we intentionally had the agent begin the session saying, "Hello, [person's name]." Three participants explicitly mentioned this feature, one stating, when asked if the agent had a personality, "Yes, respectful. It said, '[my name]', and 'I agree with this.'...I thought that was very funny. That was really cool."

Other comments implying a personality included, "Seemed a lot like a travel agent that was in a hurry," and "helpful, but kind of annoying," and "he seemed almost irritated when I didn't agree with him." One participant who did the editing task first, stated after the task that the agent did not have a personality, "It was just directed at answering questions. It had no inflections." But when asked again after the travel task, the participant responded, "It was still mechanical, but you could feel the attempt at being more personable. It acknowledged my responses, asking me to elaborate. The responses were at a more

personal level." Participants' willingness to ascribe a personality to the agent based on a few comments by the agent in one task suggests that people might be predisposed to "finding" a personality in an agent. If the effects of seeing a personality in an agent can be better understood, such a predisposition might be exploited for good purpose by designers.

Conclusion

Anthropomorphic interface agents might be one of the best interface approaches ever devised. Or they might not. Equivocal results from prior research make it virtually impossible to decide this matter. The difficulty with prior work has been its lack of systematicity in examining key factors and the use of dependent measures that often did not appropriately assess subjective experience and objective performance.

In this paper we introduced a framework for systematically examining the effects of anthropomorphic agents on user performance and subjective responses. We performed an initial experiment within this framework that suggested that type of task may play an outsized role in the perception of agents. We plan to use our framework to guide additional studies and hope other researchers find it useful and that it will allow future experiments to build on each other more effectively than in the past.

Table 2: Responses to Questionnaire Items as a Function of Type of Agent and Task.

	Animated (*n*=12)		Stiff (*n* =12)		Iconic (*n* =10)		AVG
Agent was	Travel	Edit	Travel	Edit	Travel	Edit	Travel/Edit
Worthwhile	2.50	1.58	2.25	1.42	2.30	2.10	2.35/1.57
Intrusive	2.83	3.50	3.50	4.00	3.40	3.80	3.24/3.76
Friendly	2.67	2.67	2.42	2.50	2.40	2.80	2.50/2.65
Annoying	3.25	3.33	2.83	3.25	3.20	3.80	3.09/3.44
Intelligent	2.58	2.92	2.58	2.50	2.40	2.70	2.53/2.71
Cold	3.25	3.08	3.00	2.67	3.70	3.30	3.29/3.00
Agent has clear voice	2.33	2.58	2.58	2.33	2.50	2.40	2.47/2.44
Enjoyed interacting with Agent	3.08	3.17	2.75	2.83	2.70	2.90	2.85/2.97
Agent helped with task	2.25	1.50	1.67	1.50	2.00	2.30	1.97/1.74
Like to have agent	2.83	2.67	2.58	2.33	2.20	2.40	2.56/2.47

Note: Responses were on a scale from 1 (strongly agree) to 5 (strongly disagree).

References

Cassell, J. (2000). Embodied conversational interface agents. *Communications of the ACM, 43*, 70-78.

Dahlback, N., Jonsson, A. & Ahrenberg, L. (1993). Wizard of Oz studies — why and how. In *Proceedings of the 1993 International Workshop on Intelligent User Interfaces*, (Orlando, FL), 193-200.

Dehn, D.M. & van Mulken, S. (2000). The impact of animated interface agents: A review of empirical research. *International Journal of Human-Computer Studies 52*, 1-22.

Fridlund, A.J. & Gilbert, A.N. (1985). Emotions and facial expression. *Science, 230*, 607—608.

King, W.J. & Ohya, J. (1996). The representation of agents: Anthropomorphism, agency and intelligence. In *Proceedings CHI ' 96 Conference Companion*, 289-290.

Koda, T. (1996). *Agents with faces: A study on the effect of personification of software agents*. Masters thesis, MIT Media Lab, Cambridge, MA.

Lanier, J. (1995). Agents of alienation. *Interactions 2*, 66—72.

Laurel, B. (1990). Interface agents: Metaphors with character. In B. Laurel (Ed.), *The art of human-computer interface design*. Reading, MA: Addison-Wesley, 355-365.

Lyman, P. & Varian, H. (2002). *How Much Information?* Find at http://www.sims.berkeley.edu/how-much-info/.

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM, 37*, 31-40.

McCrae, R. & Costa, P. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

Nass, C., Isbister, K. & Lee, E. (2000). Truth is beauty: Researching embodied conversational agents. In J. Cassell, S. Prevost, J. Sullivan, and E. Churchill (Eds.), *Embodied conversational agents*. Cambridge, MA: MIT Press, 374-402.

Nass, C., Steuer, J., & Tauber, E. (1994). Computers are social actors. In *Proceedings of CHI '94*, 72-78.

Rickenberg, R. & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In *Proceedings of CHI 2000*, 329-336.

Shneiderman, B. & Maes, P. (1997). Direct manipulation vs. interface agents. *Interactions, 4*, 42-61.

Shneiderman, B. (1997). Direct manipulation versus agents: Paths to predictable, controllable, and comprehensible interfaces. In J.M. Bradshaw (Ed.), *Software agents*. Cambridge, MA: MIT Press, 97-106.

Takeuchi, A. & Nagao, K. (1995). Situated facial displays: Towards social interaction. In *Proceedings of CHI ' 95*, 450-455.

Walker, J.H., Sproull, L., & Subramani, R. (1994). Using a human face in an interface. In *Proceedings of CHI ' 94*, 85-91.