

WHY PEOPLE HATE THE PAPERCLIP:
LABELS, APPEARANCE, BEHAVIOR AND
SOCIAL RESPONSES TO USER INTERFACE AGENTS

Symbolic Systems Program
Stanford University

Luke Swartz
June 12, 2003

I certify that I have read this thesis and that in my opinion it is
fully adequate, in scope and quality, as a thesis for the degree of
Bachelor of Science with Honors.

Clifford Nass
(Principal Advisor)

I certify that I have read this thesis and that in my opinion it is
fully adequate, in scope and quality, as a thesis for the degree of
Bachelor of Science with Honors.

Bonnie A. Nardi
(Second Reader)

Abstract

User interface agents are increasingly used in software products; perhaps the best-known user interface agent is the Microsoft Office Assistant (“Clippy the Paperclip”). This thesis explores why many people have a negative response to the Office Assistant, using a combination of theoretical, qualitative, and quantitative studies. Among the findings were that labels—whether internal cognitive labels or explicit system-provided labels—of user interface agents can influence users’ perceptions of those agents. Similarly, specific agent appearance (for example, whether the agent is depicted as a character or not) and behavior (for example, if it obeys standards of social etiquette, or if it tells jokes) can affect users’ responses, especially in interaction with labels.

Acknowledgements

While few of us need help from Clippy to write a letter, I needed a lot of help to write this thesis. I would like to thank the following wonderful people:

Cliff Nass—for his never-ending enthusiasm and constant support for (wow!) four years. I've had a great time working on this, and much of that is because of Cliff's great spirit.

Bonnie Nardi—for graciously agreeing to be my second reader, providing excellent advice on short notice, and for her help thinking through the interviews early on. Also thanks to her and Diane Schiano for continuing my ethnographic education this quarter.

Pamela Hinds—for introducing me to ethnography, and her indispensable advice on setting up the interviews and honing the questions

Steve Barley—for his contagious humor and down-to-earth training in ethnographic methods

Francis Lee, Amy Huang, and Young Paik, the great “I/Me Crew” —for being awesome colleagues, and for giving me the “kick in the pants” to actually do this research

Tom Wasow and the Symbolic Systems Program—for being a unique program that let me do a unique blend of research; Tom in particular has been a great mentor and friend for many years. Todd Davies also has provided excellent support and guidance for the past two years.

Laura Selznik and the Undergraduate Research Program office—for funding my research (thanks to a generous grant from the late Roger Deaton) and for their endless patience in waiting for my receipts

The informants and participants (especially the many friends who helped me pre-test the various experiments and interview questions)—for their honesty and generosity with their time; everything I learned in this research, I learned from them!

My family—for always being there, and for being my constant inspiration.

Contents

| | |
|---|-----|
| Abstract..... | iii |
| Acknowledgements..... | iv |
| Contents | v |
| Chapter 1: Introduction..... | 1 |
| 1.1 User Interface Agents | 1 |
| 1.2 Microsoft Office Assistant..... | 3 |
| 1.2.1 A Design History of the Office Assistant | 4 |
| 1.2.2 Behaviors and Roles of the Office Assistant | 5 |
| Chapter 2: Theoretical Critique of the Office Assistant | 12 |
| 2.1 Computers As Social Actors..... | 12 |
| 2.2 Critique of Agents and Anthropomorphism | 13 |
| 2.2.1 Direct Manipulation versus Agents | 18 |
| 2.2.2 The “Persona Effect” and Empirical Studies of the Effect of Using Agents..... | 19 |
| 2.3 Applying CASA..... | 21 |
| 2.3.1 Etiquette | 21 |
| 2.3.2 Appearance | 22 |
| 2.3.3 Status..... | 23 |
| Chapter 3: Qualitative Study of the Office Assistant..... | 25 |
| 3.1 Methods and Informants | 25 |
| 3.1 In-Depth Interviews | 25 |
| 3.2 Survey Question..... | 25 |
| 3.2 Results and Discussion | 26 |
| 3.2.1 General Response to the Office Assistant..... | 26 |
| 3.2.2 Expertise, Help and Learning | 28 |
| 3.2.3 Mental Model of the Paperclip | 29 |
| 3.2.4 Character Appearance..... | 30 |
| 3.3 Alternative Explanations..... | 31 |
| 3.3.1 Attitudes Toward Microsoft..... | 31 |

| | |
|--|----|
| Chapter 4: Quantitative Study of User Interface Agents | 32 |
| 4.1 Method | 32 |
| 4.1.1 Design and Manipulation..... | 32 |
| 4.1.2 Participants and Procedure..... | 33 |
| 4.1.3 Dependent Measures | 34 |
| 4.2 Results..... | 35 |
| 4.3 Discussion | 36 |
| Chapter 5: Second Quantitative Study | 39 |
| 5.1 Method | 39 |
| 5.1.1 Design and Manipulation..... | 39 |
| 5.1.2 Participants and Procedure..... | 41 |
| 5.1.3 Dependent Measures | 41 |
| 5.2 Results..... | 42 |
| 5.3 Discussion | 48 |
| Chapter 6: Conclusions | 51 |
| References..... | 52 |
| Appendix A: Quantitative Questionnaire..... | 58 |

Chapter 1: Introduction

“I hate that #@%# paperclip!” Many people seem to dislike Microsoft’s Office Assistant—why? What can one learn from the Office Assistant about how to design user interface agents? This thesis tackles these two questions. Chapter 1 provides an overview of user interface agents and the Office Assistant. Chapter 2 examines the previous literature on user interface agents to critique the Office Assistant. Chapter 3 describes a qualitative study of Microsoft Office users, and Chapters 4 and 5 describe quantitative studies designed to test design issues provoked by the qualitative study.

1.1 User Interface Agents

User interface agents are increasingly employed to enhance software products. Websites (*e.g.* buy.com, extempo.com, ananova.com, mysimon.com) now use characters to guide users through processes or present information, “wizards” and “guides” have become standard user interface tools, and a new crop of software that uses Microsoft Agent is beginning to bring anthropomorphic characters to the desktop—including a Bible-reading character (Figure 1)!



Figure 1. Agent screenshot from Speakingbible.com

While some of the hype around agents has died down, not long ago, Nicholas Negroponte (1995) predicted, “The future of computing will be 100% driven by delegating to, rather than manipulating, computers.” About the same time, Microsoft chairman Bill Gates (1995) gushed about how “the social interface” using agents would be the next step beyond the graphical user interface, making it “100 times easier to use than today’s VCR is.”

Agents also feature prominently in science fiction. Stanley Kubrick’s *2001: A Space Odyssey* depicted the interface to the computer system HAL 9000 as an intelligent agent, who not only could interact with astronauts using speech, but could play chess and even read lips. *Star Trek: The Next Generation* depicted not only a ship computer with a strong personality, but also an android character (“Data”) who behaved like a normal (albeit humorless) crewmember. The idea of artificial beings designed to assist humans is an ancient one—at least as old as Homer’s description the god Hephaestus creating golden servants to do his bidding.

What is a “user interface agent”? Unfortunately, “agent” has come to mean many things in the Computer Science and Human-Computer Interaction literature. On one hand, “agent” can describe a system designed to mimic human behavior on some level—an interpretation most associated with Artificial Intelligence (AI) and the idea of “intelligent agents.” In their definitive textbook on agent-centered AI, Russel and Norvig (2003, p. 4) use the related term “rational agent” to describe a program that “acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome.”

On the other hand, “agent” can describe software that acts on one’s behalf to carry out (relatively) independent tasks—not unlike a travel agent. They are often referred to as “autonomous agents,” or as “software agents” or, simply, “bots.” They include most “web agents,” which scour the web and report back what they’ve found.

Finally, “agent” can describe various interface elements—often involving anthropomorphic characters—designed to make a user interface more fun and/or easier to use. Such agents, which I refer to as “user interface agents,” are related to “social,” “embodied,” “conversational,” “life-like,” “animated,” and “personified” agents (or some combination of these adjectives). I prefer the term “user interface,” because it does not make a commitment to a particular representation or behavior—merely that the agent is designed to enhance the user interface.

An additional difficulty is that often an agent can fall into more than one—even all three—of these categories. For example, Apple Computer’s famous Knowledge Navigator concept video

(Dubberly & Mitsch, 1987) depicted a bow-tied agent which, while enhancing the user interface of a suped-up Macintosh-like operating system, also seemed to be intelligent (in that it understood natural language, made inferences based on the user's input, etc.) and autonomous (in that it carried out activities in the background, such as trying to connect a colleague on the videophone and leaving a message on the user's behalf).

1.2 Microsoft Office Assistant

Perhaps the most well-known user interface agent is Microsoft's Office Assistant, bundled with its Office software suite since 1997. Popularly known as "Clippy the Paperclip" (the default character, referred to in Microsoft Office itself as "Clippit"), the agent seems to have attracted widespread negative opinion. The press—particularly the digerati media—almost universally condemned the paperclip. One representative article, "Die Clippy, Die," describes how to permanently remove the persistent character (Noteboom 1998). Nearly every website about the paperclip in 2001 (before the introduction of Office XP) showed how to remove or disable it. At least among the technologically elite, Clippy was—and is—extremely unpopular.

The clamor against the character forced Microsoft to allow users to permanently hide the paperclip in the 2000 version of Office. Later, Microsoft effectively removed the Office Assistant by disabling it by default in Office XP. Microsoft claimed that the new version was so easy to use that the Assistant was no longer necessary, but it also hammed up the idea that "Clippy" was annoying by having Gilbert Gottfried's trademark screechy voice play the character during the launch. It also sponsored mocking websites (see Figure 2) encouraging users to vent their frustrations towards the paperclip—and, by extension, to buy the upgrade.



Figure 2. Image from the Microsoft homepage

Why did this seemingly innocuous character engender such vituperation? What lessons can we take from the Office Assistant when designing future user interface agents? Let us begin by describing the Office Assistant's history and behavior.

1.2.1 A Design History of the Office Assistant

The Office Assistant traces its lineage back to Microsoft Bob, a product announced in 1995 as part of the “Microsoft Home” software line. The software was inspired by Packard Bell Navigator’s “room” interface (also common to a number of Hypercard stacks and the Magic Cap system, as noted in Winograd (1996)), as well as then-recent research on social responses to computer technology, in particular, Computers As Social Actors (CASA) theory.

CASA theory, first proposed in Nass, Steuer, and Tauber (1994) and later refined and expanded in Reeves and Nass (1996), states that users instinctively treat computers like people. Since CASA showed that computers behave like social actors, it was reasoned, perhaps adding an anthropomorphic character would make a program more natural: one knows instinctively how to respond to people, so one would know instinctively how to respond to the character.

Furthermore, it was argued, if Microsoft created a popular character, it could be a commercial success in its own right: They recalled that revenues from California Raisins merchandise exceeded sales of the entire worldwide raisin industry (Cuneo, 1988). Thus, Bob (code named “Utopia”) included a number of professionally designed cartoon user interface agents, which guided the user through the program.

Bob was a commercial failure; pundits disagreed on exactly why: Was the “social interface” a failed concept, or was it merely a combination of technical difficulties (the program required a then-powerful computer to run, and even then it ran slowly) and poor marketing? In any case, Bob’s cartoon user interface agent technology was folded into the next release of Microsoft Office, Office 97 (and Office 98 for the Macintosh). This technology was combined with the Answer Wizard help query system, which had been previously deployed in Office 95

(Heckerman & Horvitz, 1998), as well as the Lumière user goal modeling system (Horvitz, Breese, Heckerman, Hovel & Rommelse, 1998), both of which used Bayesian modeling techniques.

1.2.2 Behaviors and Roles of the Office Assistant

What exactly does the Office Assistant do? As a user interface agent, it enhances the user interface; it is also somewhat autonomous (see “Proactive Help System” below) and intelligent (see “Natural Language Help Query,” below). The Office Assistant fulfills three major roles:

Proactive Help System

The Office Assistant gives help proactively—that is, it suggests ways that it can help the user finish a task better or easier. Perhaps the most famous example of this is that for writing letters: if a user types something resembling a salutation (*e.g.* “Dear John,”) into a document, the Office Assistant appears and offers help with writing the letter (Figure 3).

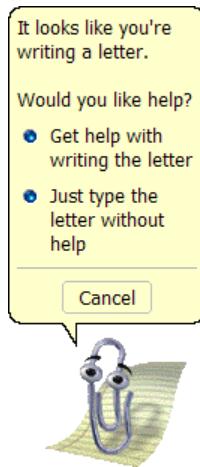


Figure 3. The infamous letter-writing proactive help feature.

Clicking “Get help with writing the letter” brings up the Letter Wizard, which aids with formatting and layout (Figure 4). Note that the Letter Wizard can also be invoked on the user’s initiative (by choosing “New...” from the “File” menu and selecting “Letter Wizard”).

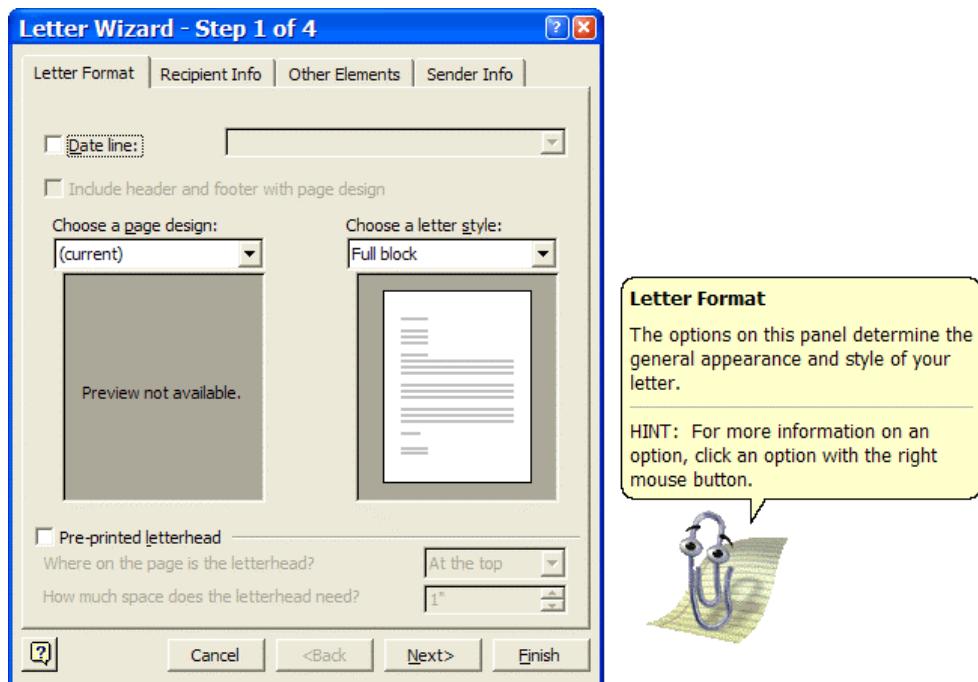


Figure 4. The Letter Wizard; what appears if one clicks “Get help with writing the letter.”

Curiously, Office Assistant proactively offers letter-writing help regardless of how many times one has clicked “Just type the letter without help.” The agent also appears even if it has been hidden (note below that *repeated* hiding in versions 2000 and above does allow one to turn off the Assistant entirely—although not this specific feature).

Similar proactive help features are “tips” triggered by user behaviors, designed to teach users about the program’s features. For example, typing a line in all uppercase and pressing return results in a tip explaining the Headings feature (Figure 5).

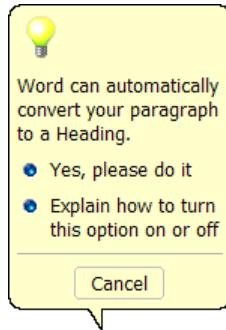


Figure 5. Proactive help tip on Headings.

Similarly, some tips explain features when triggered by signs that the user might be “struggling.” For example, clicking repeatedly in the margin (where one generally can’t type) causes the Office Assistant to display a light bulb, which, when clicked, explains how to enter text in that area.

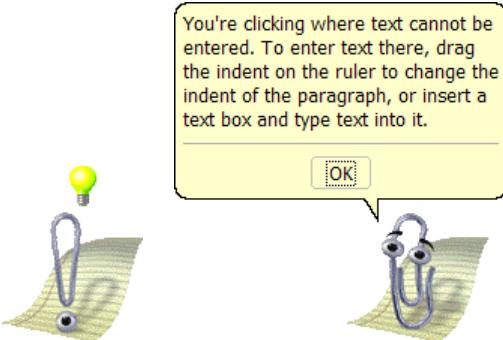


Figure 6. Proactive help tip triggered by margin clicks, before and after the user clicks the light bulb

One level of proactivity below the “light bulb” are tips shown only when the user clicks on the agent. For example, if the user has been working with tables and then clicks on the Assistant, the Assistant will guess that the user may want to know more about certain table features, and will show a selection of them (Figure 7). The help is still “proactive” in the sense that the agent proactively suggests helpful tips, even though users only see the suggestions if they click on the

Office Assistant. This feature is perhaps the closest to the Microsoft prototype Lumière Project described in Horvitz *et al.* (1998).

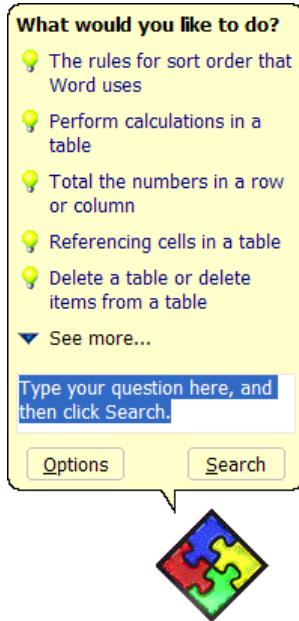


Figure 7. Tips displayed after the user clicks on the Office Assistant.

Unlike the letter-writing proactive help feature, tips only appear once, and they do not appear unless the Office Assistant is already visible. Users can also configure the Office Assistant to display a random tip when the program starts—allowing the agent to be purely pedagogical (rather than proactively helping with a specific need).

Natural Language Help Query

As noted in section 1.2.1, the Office Assistant incorporates a version of the Answer Wizard feature, developed initially for Office 95 (Heckerman & Horvitz, 1998). The Answer Wizard uses basic Bayesian inference to guess a users' goal, given a particular help query. This allows users to ask questions in relatively natural language, and generally provides better results than a mere “keyword” search. (Note that the original Answer Wizard interface, which does not use an animated character, remains in Office, and can be invoked if the Office Assistant is turned off.)

Upon submitting a query, the Office Assistant provides a list of help topics that it thinks would be useful (see Figure 8); each link displays help on a particular topic in a separate help window.

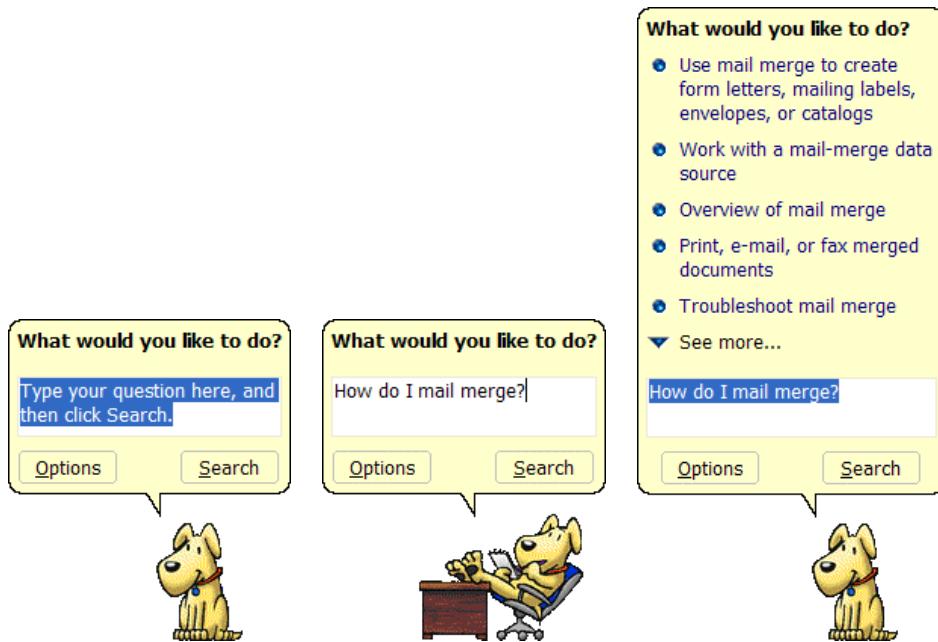


Figure 8. Natural language help query; the user clicks on the Office Assistant, types a question, and clicks Search.

Note that the help displayed is the same as the help available without the Assistant. That is, if one turns off the Assistant, the Answer Wizard provides the same help choices, which lead to the same help screens.

Agenthood of the Program

Finally, the Office Assistant acts as a locus of agenthood for the program. For example, it is the “voice” for dialog alert boxes (Figure 9).



Figure 9. A dialog alert box “voiced” by the Office Assistant

Likewise, certain commands (saving or printing, sending an email) cause the Assistant to display an animation of that action, suggesting that it is, at least in part, responsible for that action. Finally, the agent does small “social idle” animations (blinking, etc.) when on-screen and nothing is going on (i.e. the user isn’t typing or clicking), a feature developed for Bob to make the characters more amusing and natural (Winograd, 1996, p. 149).

Office 2000 adds the option of turning off the Assistant entirely (as noted earlier, this is the default in Office XP). The newer versions also allow one to turn each of the roles above on and off (see Figure 10). For example, unchecking “Display alerts” disables the dialog box “agenthood” feature, such that the program returns to displaying standard dialog boxes regardless of whether the Assistant is on screen or not.

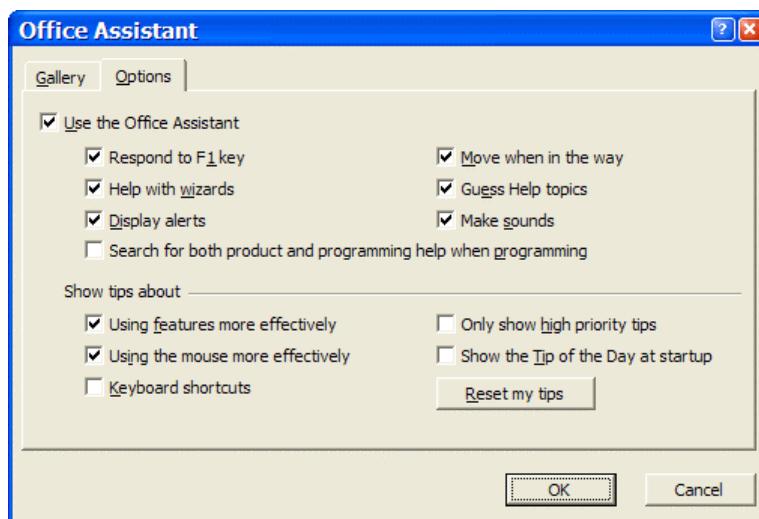


Figure 10. Options for Office Assistant (from Office 2000)

Office 2000 also adds an automatic disabling feature: If one hides the Assistant several times, it responds with Figure 11, an option to turn it off “permanently” (although this merely unchecks the “Use the Office Assistant” box, and is thus not necessarily permanent).

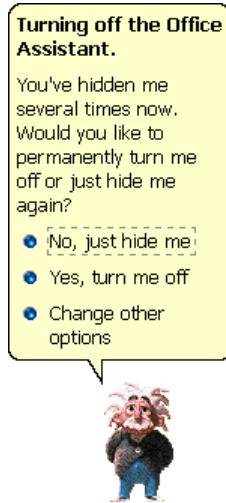


Figure 11. Automatic disabling (from Office 2000)

Chapter 2: Theoretical Critique of the Office Assistant

What findings in the existing Human-Computer Interaction literature can help explain why people dislike the Office Assistant? Unfortunately, despite the Office Assistant's ubiquity and notoriety, it has received little or no research attention. Nevertheless, there are several theories and debates regarding user interface agents that are relevant to the Assistant.

2.1 Computers As Social Actors

As noted in Chapter 1, the Office Assistant was in part inspired by the Computers As Social Actors (CASA) theory that people instinctively treat computers and other media as if they were real people (Nass, Steuer & Tauber, 1994; Reeves & Nass, 1996).

The theory is based on a large body of research showing how human psychological phenomena can be replicated on a computer. For example, studies found that people subconsciously use the same standards of politeness (Nass, Moon, & Carney, 1999), gender stereotypes (Nass, Moon, & Green, 1997), teamwork (Nass, Fogg, & Moon, 1996), and reciprocity (Fogg & Nass, 1997) in their interactions with computers as they use with other people. Similarly, there is evidence that people will rate computers with similar personalities to themselves higher—just as people will rate other people with similar personalities higher (Nass, Moon, Fogg, Reeves, & Dryer, 1995; Moon, 1998; Moon & Nass, 1996; Moon & Nass, 1998; Nass & Lee, 2000).

It should be stressed that CASA theory refers to *unconscious* social responses—in fact, many of the participants who are questioned after experiments emphatically deny the very behavior they just exhibited (Reeves & Nass, 1996). Reeves and Nass speculate that humans evolved to assume that objects exhibiting certain human-like traits are actually human. Thus, modern humans presented with interactive media will unconsciously respond to those media in a social way—even if they *know* (consciously) that those media are in fact not real humans.

What kinds of traits trigger this unconscious response? Nass and Moon (2000) note that exactly what interfaces will trigger what social responses remains a largely unsolved question.

Nevertheless, Nass and Steuer (1993) suggest four characteristics that each “strongly cue the idea that one is interacting with a social actor”: language use, interactivity (defined as how much the system uses prior input to determine its subsequent behavior), playing a social role (*e.g.* doctor, tutor, parent), and having human-sounding speech. It is also clear that, as Reeves and Nass (1996, p. 25) point out, “people don’t need much of a cue to respond socially”—indeed, plain text has been found to trigger the same social responses as more vivid technologies such as voice (*Ibid.*). In fact, Steuer (1994) found that a text-based tutor was perceived as *more* “like the user” and likeable than a full-motion video tutor. Nass, Steuer, Henrickson, and Dryer (1994) also emphasize that “minimal social cues can induce computer-literate individuals to use social rules.”

These two aspects of CASA theory—its unconsciousness and the relatively simple ways to trigger social reactions—are often forgotten when applying the theory to design. While the CASA studies show that people unconsciously respond to even simple computer interfaces in social ways, they do not (necessarily) show that people will like or benefit from computer interfaces which consciously try to behave as social agents. It is one thing to take advantage of unconscious social responses, and quite another to make that response explicit by displaying an anthropomorphic character that asserts its agenthood.

This distinction implies that the Office Assistant, while inspired by CASA findings, is not in itself justified by those findings. Nevertheless, a number of authors (*e.g.* Laurel, 1990) suggest that because most computer use is unavoidably social, explicitly social agents make the interface easier to learn and use, because people naturally know how interact socially.

2.2 Critique of Agents and Anthropomorphism

“Never trust anything that can think for itself if you can’t see where it keeps its brain”
— *Harry Potter and the Chamber of Secrets*, p. 329.

Despite enthusiasm from CASA-inspired designers, some authors have criticized the idea of explicitly social, anthropomorphic agents. For example, Ben Shneiderman (1989) refers to anthropomorphism as the “humpty dumpty syndrome” and implores designers to resist being “seduced” by this “primitive urge.” Jaron Lanier (1996) puts it succinctly: “Intelligent agents stink.” (He has also referred to the idea of such agents as “wrong and evil” (1995).)

Some of the critiques of agents and anthropomorphism are rather easily dismissed:

Lazy Programmers and Quirky Interfaces

Lanier (1996) claims that the “autonomy” provided by agents makes programmers lazy “because then [the program] has the right to be quirky.” While this is a legitimate worry—programmers and designers may indeed accept unreasonably unpredictable interfaces when designing agents—this need not apply to all agents. Indeed, Laurel (1991, p. 145) claims that, in fact, caricatured “dramatic characters are better suited to the roles of agents than full-blown simulated personalities.” That is, the best-designed agents are more predictable than people, because they rely on stock dramatic archetypes rather than “quirky,” idiosyncratic personalities.

Annoying and Distracting Characters

Shneiderman (1995) claims, “The anthropomorphic styles are cute the first time, silly the second time, and an annoying distraction the third time.” Again, this is a legitimate worry, especially for user testing—if one merely tests an interface once, annoyances may not present themselves. However, whether a character will be annoying or not largely depends on its behavior: If the anthropomorphic agent almost always presented useful information in an easy-to-understand way, perhaps it would not be annoying or distracting. Indeed, some agents are being designed to *reduce* the number of distractions one must endure (Markoff, 2000). Moreover, as we will discuss in Chapter 3, some users *prefer* occasional distraction because it is entertaining.

Poor Market Performance

Another argument is that various anthropomorphic interfaces, such as Postal Buddy, Microsoft Bob (Shneiderman & Maes, 1997), and anthropomorphic bank terminals such as Tellie the Teller, Harvey Wallbanker, and BOB The Bank of Baltimore (Shneiderman, 1995) have all

failed in the marketplace. While one should learn from this history, one cannot infer that all anthropomorphic interfaces are doomed to failure based solely on these products' market performance. For instance, it is possible that these failures depended on the specific implementation (*e.g.* Bob's slowness) or the domain (*e.g.* ATMs may not be an appropriate place for agents). Furthermore, Microsoft Office was *not* a poor market performer—although one might argue that this was in spite of, not because of, the Office Assistant feature. Either way, it offers a useful lesson: one cannot rely solely on market data to determine whether a particular feature is good or not.

Other criticisms of agents and anthropomorphism, however, are more substantive and deserve closer study:

Anthropomorphic Dissonance: Misleading Expectations and Erroneous Conceptual Models

As early as 1980, Sheiderman noted that natural language technology can lead to “unrealistic expectations of the computer’s power” (p. 208). Similarly, he argues more recently, appearances and behaviors that attribute autonomy to a computer “can deceive, confuse, and mislead users” (Shneiderman, 1998). Watt (1998) refers to the gap between expected behavior given an agent’s appearance and its actual behavior as “anthropomorphic dissonance”: “The bigger the gap, the greater the dissatisfaction with the interface.” That is, presented with an agent that putatively understands natural language or acts “like a person” (encouraged either by marketing or its appearance to believe such things), users will expect more from the interface than it is capable of, leading to inevitable disappointment, frustration, and dissatisfaction. It is possible that the Office Assistant, by presenting itself as having natural language facility, encourages these over-optimistic expectations. Its anthropomorphic appearance might also lead users to over-estimate its abilities as compared with the Answer Wizard, which has the same natural language technologies but no animated character.

Moreover, argues Shneiderman (1989), such agents and characters can cause people to form “an erroneous model of how computers work and what their capacities are.” Perhaps anthropomorphic interfaces encourage people to think about computers in ways that do not reflect how they actually work, thus making using the computer more difficult because of the

flawed conceptual model. Chapter 3 discusses how many novice users' conceptual models of the Office Assistant were indeed flawed or confused—perhaps encouraged by its anthropomorphic character. However, one must wonder whether a properly presented agent might result in more accurate, useful conceptual models. For example, Shneiderman (1998) cites Resnick and Lammers (1985) to show that “[s]ubjects reported being less confused” when given “constructive” (what the authors refer to as “neutral”) than with “human-like” or “condemning” (what the authors refer to as “computer-like”) feedback. However, the confusion may have been inherent in the manipulation: “I don’t understand these numbers” hardly means the same thing as “Use letters only.” Perhaps there is a way to communicate a clear conceptual model of the agent while still using human-like dialog (*e.g.* “I’d appreciate it if you would only use letters” or “I only understand letters”).

While anthropomorphic dissonance presents a real pitfall for agent design, critics sometimes take this argument too far, claiming that anthropomorphic agents will blur the line between humans and computers. Shneiderman (1989) suggests that this is especially important for children, since “it is important for children to have a clear sense of their own humanity.” Lanier (1995), in view that “there is nothing more important to us than our definition of what a person is,” claims that agents “make people diminish themselves” and “redefine themselves into lesser beings.” Is it true that people cannot distinguish an anthropomorphic computer from a real human being? Laurel (1991, p. 143) argues that “[t]here is no evidence to suggest that computer-based characters, no matter what the degree of lifelikeness, lead people to believe that either the machine or the characters themselves are actually alive.” Indeed, there does not seem to be any substantive evidence that people consciously believe that computers (or anthropomorphic agents represented on computers) are actually people. The philosophical debate over what constitutes humanity is indeed an interesting and important one, but it is beyond the scope of this thesis—and likely does not affect people’s actual responses to user interface agents.

The Computer Did It: Lower Feelings of Control and Self-Reliance

Shneiderman claims that agents can reduce users’ feelings of self-reliance and control: “I think anthropomorphic representations destroy the users’ sense of accomplishment; I think users want to have the feeling they did the job—not some magical agent” (Shneiderman & Maes, 1997).

Instead, he argues, one should work to create interfaces encouraging an “internal locus of control” so that users feel that the computer is a tool over which they have mastery, rather than an autonomous agent (Shneiderman, 1989; Shneiderman, 1998).

Shneiderman cites two studies to support this claim: Quintanar, Crowell, and Pryor (1982) showed that while students scored higher using an anthropomorphic interface, they felt less responsible for their performance. Gay and Lindwarm (1985) showed that, given an interface using anthropomorphized prompts (*e.g.* “Hi! I am the computer. I am going to ask you some questions.”) or neutral prompts (*e.g.* “This is a multiple-choice exercise.”), users were more likely to change their opinions to think that computers are harder to use, as compared with those given an interface using “you” prompts (*e.g.* “You will be answering some questions.”). Thus, in both studies, users in anthropomorphic conditions felt a loss of control.

While agent proponents might argue that agents necessarily involve giving up control in exchange for substantive benefits (such as not having to worry about a task), when designing help agents such as the Office Assistant, it would seem important to preserve and expand users’ sense of control and self-reliance. Perhaps a less-anthropomorphic version would make the user feel more in control. In fact, Office XP includes a new “task panes” feature, which duplicates some of the Office Assistant functionalities without using an animated character: perhaps such non-anthropomorphic panes lead to more feelings of self-control. However, it seems at least possible to design an agent that actually *increases* users’ feelings of control and self-reliance: for example, agents could teach users a skill, empowering them to finish a task on their own the next time. To some degree, this is highlighted in the difference between the Office Assistant’s letter-writing proactive help feature and most other “tip” help features: the letter-writing feature constantly asks users if they want help (creating or at least suggesting a dependent relationship), while the tips teach users a skill once, leaving them to use that skill on their own the next time (and thus increasing the users’ feeling of control).

2.2.1 Direct Manipulation versus Agents

Coined by Shneiderman (1983), “direct manipulation” refers to a method of controlling a computer by directly manipulating interface elements—such as dragging an item around the screen. Some kinds of agents—particularly autonomous ones—seem opposed to this metaphor, as they work using indirect delegation and management rather than direct manipulation. For some direct manipulation enthusiasts, thus, agents represent a step backwards in user interface technology, returning to something akin to pre-GUI command-line dialog interfaces.

This has led some people to see the issue as a “debate” between direct manipulation and agents (*e.g.* Shneiderman & Maes, 1997). One might argue that this debate was touched off by Maes’ oft-cited paper on agents (1994), which claims that the direct manipulation “metaphor will have to change if untrained users are to make effective use of the computer and networks of tomorrow.”

Of course, it need not be a simplistic, either-or decision: In response to the frustrated criticism, “Why should I have to negotiate with some little dip in a bow tie when I know exactly what I want to do?”, Laurel (1990, p 356-357) responds that some tasks are more suited to agents than others: “Few of us would hire an agent to push the buttons on our calculator; most of us would hire an agent to scan 5,000 pieces of junk mail.” Maes also notes that some tasks are better delegated to agents than others—for example, she admits that while having someone else fix her car means that she gives up both control and understanding of how the car works, she nevertheless prefers delegating the task (Shneiderman & Maes, 1997).

Horvitz (1999) suggests that, rather than choose between direct manipulation and “automation,” one can seek “valuable synergies” between the two interface techniques in a mixed-initiative system. Susan Brennan (1990) even argues that agent-like conversation can be a form of direct manipulation—and that direct manipulation succeeds partly because it shares features with conversation. Similarly, Rickenberg and Reeves (2000) argue that trying to generalize agents as good or bad is like trying to generalize film or the internet as good or bad: there are a number of both positive and negative factors, each of which depend largely on the situation.

Is an agent appropriate to Microsoft Office’s situation? Doyle (1999) argues that the Office Assistant is not badly designed but rather “chosen for the wrong domain.” He notes that “[b]uilding a spreadsheet, for example, is essentially a mechanical task—entering numbers and equations—and not one about which a user is likely to want discussion.” However, this argument has two problems: First, as we’ll discuss in section 2.3, the Microsoft Office Assistant does have some serious design problems. Second, the argument that spreadsheets are mechanical ignores research (*e.g.* Nardi & Miller, 1990) showing that spreadsheets actually act as collaborative “cognitive artifacts,” about which there is a good deal of discussion. It’s true that the actual data entry is fairly mechanical and not well-suited to adding an agent, but there may be other tasks involving spreadsheets for which agents might prove useful. (Indeed, one might consider Microsoft Excel’s AutoFill feature to be a primitive autonomous agent, showing how agents can be useful for even data entry!)

While we might not throw out the idea of an agent being useful to office productivity software in general, all these authors have a point: some situations and tasks are more suited to using agents than others.

2.2.2 The “Persona Effect” and Empirical Studies of the Effect of Using Agents

What research, then, has been done to determine the positive and negative effects of agents in various situations?

Dehn and van Mulken (2000) present an excellent overview of the bulk of the empirical research that has compared various user interface agents to interfaces without such agents. They note that much of the research is plagued by methodological problems: For example, Lester, Stone, Converse, Kahler, and Barlow (1997) vary not only the presence of an agent but also what advice is given. Similarly, Sproull, Subramani, Kiesler, Walker, and Waters (1996) vary the presence of an agent, but the agent presentation also introduces a one-second delay between the appearance of the agent and hearing the text.

Sproull *et al.*'s agent was a three-dimensionally modeled face with an unnatural voice without inflection; several studies at Stanford (*e.g.* Flannery & Merrill, 2000) suggest that some three-dimensionally modeled faces (such as the “Baldi” character from the CSLU Toolkit; see Cole *et al.*, 1999) are perceived as being “weird” and thus serve as a cognitive distractor. McBreen, Shade, Jack, and Wyard (2000) also found that three-dimensional talking heads were perceived badly (compared to video, disembodied voice, and still images), partially because of bad lip synchronization.

Despite these methodological difficulties, Dehn and van Mulken attempt to draw some broad conclusions from the research. First, the “persona effect” posited in Lester *et al.* (1997)—that characters positively influence users’ responses to a system—found some support: in particular, it seems clear that adding agents tends to make the system more entertaining. Other effects, they argue, depend on “what particular anthropomorphization is chosen and...the domain in which the interaction is set.”

These trends also bear themselves out in studies not included in Dehn and van Mulken’s meta-analysis: For example, Moundriou and Virvou (2002) found that, while instructional agent conditions didn’t result in better learning than a non-agent control condition, the system was rated as being more enjoyable and problems were perceived as being less difficult. Wexelblat (1998) also found that anthropomorphic interfaces were rated as being more enjoyable and likeable. Likewise, Dehn and van Mulken’s caution that certain kinds of agents fare better than others is borne out in other research: Lee and Nass (1999), like McBreen *et al.* (2000), found that sometimes less “vivid” representations of an agent—such as a text box or disembodied voice—were rated higher than poorly implemented representations (such as a stick figure and a three-dimensionally modeled talking head).

How can one apply this research to the Office Assistant? It seems unclear whether an agent—and, specifically, an anthropomorphic character agent—is appropriate to the office context. However, it seems clear that such an anthropomorphic character agent, if properly designed, could make the program more entertaining, enjoyable and likeable. We shall next question why many users find the Office Assistant annoying—and not enjoyable.

2.3 Applying CASA

While CASA theory does not, in itself, justify the use of user interface agents, anthropomorphic or not, if one *does* use an anthropomorphic interface, that will certainly be enough to trigger an unconscious—and, likely, conscious—social response to the agent. Thus, the CASA paradigm can be applied to analyze the Office Assistant itself: If people (consciously or unconsciously) treat the Assistant like a person, then how can we predict and explain their responses to it?

CASA theory predicts that psychological rules that apply to people will also apply to interactions with a computer (or, in this case, an agent). So, one must ask, “What would one want in a human assistant?” Many of the Office Assistant’s behaviors would be outright intolerable in a human, such as continuing to ask the same question over and over.

The Office Assistant also breaks more subtle rules of human-human interaction, such as staring at the user and monitoring the user’s work. For example, Rickenberg and Reeves (2000) found that people who performed a task while an agent monitored them had both higher reported anxiety and lower performance on the task. Perhaps a more successful, less anxiety-creating, Office Assistant would have a desk of its own to work at, minimize itself into an unobtrusive icon, or even turn away from the user when not called into service.

2.3.1 Etiquette

Rules of human social interaction can be grouped under the name “etiquette,” an increasingly important sub-field of user interface agent research. Bickmore (2002) describe etiquette as “adhering to prescribed norms in social interactions, or about negotiating and making explicit interactional norms when they do not already exist.” To that end, Miller and Funk (2001) propose a short list of etiquette “rules,” such as “Don’t make the same mistake twice.” Since the Office Assistant continues to persist in displaying its letter-writing proactive help feature despite being dismissed an arbitrary number of times, it obviously breaks this rule! Interestingly, Miller and Funk also encourage agents to “talk explicitly about what you’re doing and why” in a sort of

“meta-communication”; thus, users can be well-informed and develop a more accurate conceptual model of the agent.

Bickmore and Cassell (2001) note also that both Microsoft Bob and the Office Assistant use “essentially a passive strategy for relationship building.” They claim that human etiquette rules demand an active strategy for “building, maintaining or changing a relationship with the user”: Just as a human assistant would learn one’s preferences over time, so to should a user interface agent develop a rich, long-term user model.

The Office Assistant’s letter-writing proactive help feature, thus, breaks every relevant etiquette rule: it ignores social conventions of when to disturb someone, it does not learn from its mistakes, it does not develop a long-term relationship, and (one might argue) it does not even provide a helpful service! Since this feature is the most cited annoyance in the popular press, one cannot help but wonder how much better the Office Assistant would be perceived if this *one feature* had been fixed or eliminated before its release.

2.3.2 Appearance

CASA also suggests that characters that are popular and likeable in the “real world” would be more popular and likeable on the computer. For example, animators have exploited Konrad Lorenz’s “*Kindchenschema*” for years, noting that characters with certain baby-like biological triggers, such as large heads, short arms and legs, round skulls, big eyes, and round cheeks are perceived as “cute” and likeable. Presumably, these traits would also be desirable in a likeable user interface agent character. However, the default Office Assistant character, “Clippit the Paperclip,” has virtually none of these features; especially in its original version, it features a small (or nonexistent) head, long wire “arms,” and slanted eyes (Figure 12).



Figure 12. Clippit (original version from Office 97)

Interestingly, in both static and animated tests prior to Office’s release (one of which took place in three different countries), several other characters were rated above the paperclip; the paperclip was chosen to be the default mostly because it was associated with an “office” (Nass & Reeves, 1996). Combined with *Kindschenschema* data, this suggests that Microsoft made a poor choice in selecting the paperclip as its default assistant character.

2.3.3 Status

Sara Neff (2002, pp. 29-32) has suggested applying the concept of status, an especially important one in improvisation theatre, to design. Products, she argues, each assert a different kind of status, and in turn affect the status of their users. “It is a universal rule that everybody like his or status raised,” so user-raising products will generally be more satisfying. In a way, this attribution is in line with CASA theory: just as people like other people who raise their status, they will also like products (and, presumably, user interface agents) that raise their status.

For example, Neff notes that Pull-Up diapers raise the status of toddlers, making them feel older and giving them a feeling of competence and control (their commercial jingle was “I’m a big kid now”). Similarly, Apple Computer’s Macintosh raised the status of “the rest of us”—users who were bewildered by the command-line DOS interfaces. In contrast, Neff gives the example of the VCR as a status-lowering object, as it “is in effect saying to its user, ‘You are too stupid to understand how to use me.’”

The Office Assistant *could* help raise the status of beginners, as it would provide a help function close at hand at any time—without needing to appeal to someone else. However, it can also

lower beginners' status: For example, a friend told me that she doesn't like the Office Assistant because "it reminds me of how much I don't know." For her, the Office Assistant is not unlike the flashing "12:00" on so many VCRs.

Regardless of whether the Office Assistant raises or lowers beginners' status, it would seem to lower the status of more advanced users. Lanier (1995) calls Microsoft Bob "offensively paternal"—in essence, he thinks it is status-lowering. Many advanced users don't need—or at least think they don't need—an ever-present help agent, and thus they may perceive the Office Assistant as trying to lower their status, particularly when encountering its proactive help feature.

Chapter 3: Qualitative Study of the Office Assistant

To gain a holistic view of how people use and view the Office Assistant in context, I conducted a qualitative study of Microsoft Office users.

3.1 Methods and Informants

3.1 In-Depth Interviews

In-depth interviews were conducted with 14 informants. The interviews were open-ended, took place in the space where the informants used Microsoft Office the most, and included informant-directed observation of work practices and artifacts (*e.g.* showing the last few documents used in Word). The informants were found using a variety of means, mostly recruited at a college campus over email with the enticement of a free Jamba Juice gift certificate. Because of the location in which they were recruited, most of the informants work in the education industry. However, special care was taken to recruit administrative support staff and others in less “academic” roles: No professors and only two current students were part of the informant pool.

Informants range in age from early 20’s to late 50’s, with a good distribution throughout the scale. Four were male, and ten were female; racially, nine were of European descent, three were of Asian descent, and two were of Hispanic descent. Ten informants used Microsoft Office on the PC; one used Microsoft Office on the Mac, and three used Microsoft Office on both a Mac and a PC. All names mentioned are pseudonyms; in some cases, minor details about the informants are changed to protect their anonymity. All quotes, however, are verbatim.

3.2 Survey Question

To round out the interview data, the subjects in the quantitative experiments (described in Chapters 4 and 5) were asked, “What are your (brief) thoughts on the Microsoft Office Assistant (Clippy the Paperclip)?” at the end of an online questionnaire. See Chapters 4 and 5 for more information about how participants were recruited and the nature of the pre-questionnaire task.

Their responses were coded and analyzed in a similar fashion to the interview data.

3.2 Results and Discussion

3.2.1 General Response to the Office Assistant

Informants were asked what they thought about the Office Assistant, and asked what words they associate with it. Responses varied from “I hate that guy!” to “I love it—it makes me laugh!” Half of the informants—seven—had an unqualified negative reaction to the Office Assistant. Two of the informants had unqualified positive reactions to the Assistant; the remaining five informants either had mixed responses or were confused about the Assistant (see 3.2.3).

Cute

Seven of the informants described the paperclip character as “cute,” although this wasn’t always a positive attribute. Three thought it was excessively cute; as one informant put it, “Maybe I’m just old school, but it’s too cute for me.” Another informant noted that while it is “cute and entertaining...it’s annoying when you have to work.” Indeed, five informants used the word “annoying” to describe the paperclip character.

Unnecessary

Three informants noted that one doesn’t need the Office Assistant to get help—that is, that the already-existing help features can do the same thing as its search box. They called the paperclip “stupid...needless” and “unnecessary.” Indeed, it is possible to get the same “natural language” help query via the “Answer Wizard”—whose content is taken from the same place as the “Contents” and “Index” help features.

In the Way

Four informants complained about the Office Assistant getting in their way. “It takes up space,” noted one; “It always seemed to be in the way,” said another. It is interesting that, even with the advances in Office 2000 (the character appearing on its own, outside of a window, and

automatically moving out of the way), the character still obscures one's visual field and serves as an impediment to working.

Popping Up

Three informants noted that they didn't like the proactive help feature. In the words of one informant, "I don't want it to think you need help...I want to ask for it." On reflection, one informant noted that what she finds most annoying about the Office Assistant is that "the computer is doing something you haven't told it to" and that "it challenges our authority."

Good for Other People

Six people noted that it would be useful for beginners, but not themselves. As one informant put it, "It's good for a small group of people, like my mom, who are scared of the computer...otherwise, it's patronizing." However, only one of these six recalled that the Office Assistant had actually been useful when she was a beginner; she noted that it was good to be able to "type into this box" at any point that she needed help, although now "it's a hand-holding thing" that she no longer needs.

Cognitive Labels

How can one explain the very different responses to the Office Assistant? Both those who liked and disliked the Assistant noted the qualities listed above. However, those who liked the Assistant more saw the character's animated motions as being a welcome diversion from the tedium of office work. One informant talked at length about how it would amuse her during long work periods, to the point where a friend "called it my boyfriend, since it winks at us." That is, both groups saw the Office Assistant as distracting; they differ in whether they saw this distraction as positive or negative. One can explain the dichotomy between those who liked the Office Assistant and those who did not by appealing to the *cognitive labels* they ascribed to the character. Those who labeled the character as a "productivity tool" which was supposed to be "useful" thought that its distracting animations were counter-productive and annoying—that is, trying to be "too cute." Those who labeled the character as an "office diversion" which was supposed to be "fun," by contrast, welcomed the distracting animations.

3.2.2 Expertise, Help and Learning

Informants were asked to rate, in their own words, their familiarity and history with Microsoft Office (especially Word) and word processing on a computer. Two were very clearly beginners; as one informant put it, “on a scale of one to ten, I’m a one!” Eight rated themselves as “advanced beginner[s]” or of “intermediate” expertise. Four rated themselves as being “advanced” or having “high” expertise.

Informants were also asked to characterize what they do when they need help with Office—in particular, what they did the last time they needed help—and how they learn about new features. Interestingly, answers to this question correlated with informants’ description of their expertise. Intermediate and high-expertise respondents noted that while they often learn about new features from colleagues, when they need help, they first consult the Microsoft Office help files or the Internet. Blikstein (2000) likewise found that people report learning about new features mostly from other people. Three people said they look at manuals, although two of them noted that many modern software products don’t have manuals anymore. If these methods fail, the informants said they would ask around, or perhaps post a question to an email list. The beginners, by contrast, all said that they go straight to other people for help. Take the following example:

Interviewer: Think back to the last time you needed help. What happened?

Informant: I grabbed [name] next door.

Interviewer: Where do you go for help if he’s not around?

Informant: I go to [name].

Interviewer: And if she’s not around?

Informant: I’d ask someone in [location].

Interviewer: What if nobody is around?

Informant: I’d wait!

On reflection, this behavior makes perfect sense: If one is a novice user, one is surrounded by more-experienced people. Thus, unless there is some impediment to asking help from others, it would make most sense to ask one’s co-workers for help when one runs into technological difficulties. Nardi and Miller (1990) likewise found that spreadsheets tend to be used by more than one person—the result of a collaborative effort, in which co-workers taught each other and

“subcontracted” work to each other. If the spreadsheets, reports, letters, and other documents created by Office Assistant are created in a collaborative environment, that environment is likely to be a more natural place for beginning users to get assistance than an online help tool. This suggests that the Office Assistant’s help features are largely unnecessary for novices.

3.2.3 Mental Model of the Paperclip

Informants were asked a number of questions designed to determine their mental model of the paperclip—that is, how they think it works. These questions included, “When has it appeared?”, “What do you think triggers it?”, and “What is it supposed to do?

Again, answers correlated largely with level of computing experience. Four informants—two beginners and two “advanced beginners”—seemed confused about what the Office Assistant does. One advanced-beginner informant noted that it “tells me I’ve done something wrong... It’s supposed to stop you so you don’t continue on to make a mistake.” The other confused advanced beginner said, similarly, “It tells me when I need help.” While the proactive feature does indeed try to step in when the user is attempting to do something that is impossible, this doesn’t seem to characterize the Assistant’s intended or actual role. The two beginner informants were confused as to what the paperclip did. One noted, “I don’t know what the h*** it was for. There’s no manual that tells you what it does.... The only thing I’m sure it does is it wiggles when the computer’s working.”

The other informants had more accurate mental models of the Office Assistant. They all spoke about being able to type words or questions into its search box. Two people noted that it tends to pop up when one is encountering an unfamiliar feature: “It seems to know when I haven’t done something before.” Three informants noted that it offers assistance in writing letters. Two informants associated the Assistant with other automatic tools in Microsoft Word, like AutoComplete and AutoFormat. As one put it, “it puts bullets where it thinks the bullets should be.”

Two interesting points present themselves here: First, beginners—the people who are supposed to be helped the most by the Office Assistant—are at least somewhat confused about what it is supposed to do. Especially given that beginners won't naturally turn to the computer for help (as they seek out people instead, as described in 3.2.2), it may be especially important to introduce such users to what the Assistant does and how to use it effectively.

Second, that even relatively experienced users attribute a number of actions (such as automatic formatting) to the Office Assistant suggests that users are so used to the direct-manipulation application-as-tool metaphor, that any amount of independent action will be ascribed to the agent. For these users, the agent has taken on agency for the program itself!

3.2.4 Character Appearance

Two informants said that the paperclip character “looks stupid.” Only two informants had changed the character at any time—both to the cat. “I have a cat,” one explained. To get a feel for people’s reactions to different characters, each informant was presented with images of the various Office 2000 characters, as well as “Peedy the Parrot” from Microsoft Agent, in a printout (see Figure 13).



Figure 13. Alternative characters presented to informants

None of the informants had particularly strong reactions to any character; three preferred the dog and cat characters, while one preferred the Office logo because ‘It has no eyes...it’s not

sentient.” While it seems clear that Microsoft chose a relatively unpopular look for its character, it seems that the strongest user responses are unrelated to the paperclip character itself. (Indeed, three of the informants use a Macintosh, where the default character is a classic Macintosh box with feet, not the paperclip.)

3.3 Alternative Explanations

3.3.1 Attitudes Toward Microsoft

It is possible that people with negative reactions to the Office Assistant actually have negative reactions towards Microsoft or Microsoft products, and use the Assistant as a convenient proxy upon which they project their feelings. Thus, informants were asked about how they felt about Microsoft and Microsoft products. While two informants mentioned that Microsoft is a “monopoly” and six had somewhat negative feelings about the company, negative feelings about Microsoft did not seem to correlate strongly with feelings toward the Office Assistant. For example, one informant who called Microsoft and its products “great” said that the paperclip is “annoying...I don’t like it,” while another informant who was very positive towards the paperclip noted that she doesn’t like Microsoft’s monopolistic practices.

Chapter 4: Quantitative Study of User Interface Agents

The qualitative, ethnographic study (Chapter 3) revealed a number of interesting insights into how users respond to the Microsoft Office Assistant. One surprisingly important insight was that informants' cognitive label profoundly influenced the way they perceived the agent. Is this true of other agents in other contexts—and can such an effect be shown in a controlled, quantitative study? We set out to test exactly that, varying the explicit label given to users to describe an online user interface agent. Given the interesting comments about agents' appearances, we also tested two different kinds of characters.

4.1 Method

4.1.1 Design and Manipulation

The experiment was a 2 (label: “fun” or “useful”) by 2 (character: human or cartoon) balanced, between-subjects design. The first independent variable, *label*, corresponded with how the online character was introduced. Those in the “useful” conditions were told on an introductory webpage that the character “is designed to provide useful information and make it easier to find what you’re looking for. This should make the site easier to use and help you to get your task done.” Those in the “fun” conditions, by contrast, were told that the character “is designed to make your visit more fun and give you a different way to do things. This should liven up the site and make your experience more entertaining.” The characters reinforced the label by repeating an excerpt of this text during their self-introductions (see Figure 14).

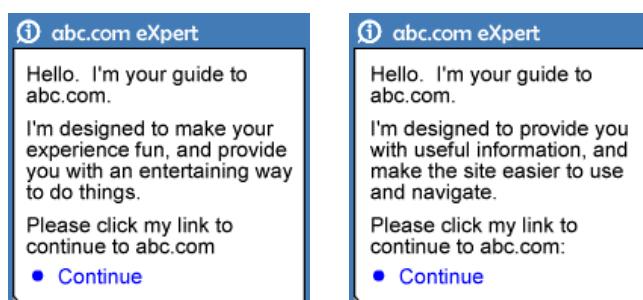


Figure 14. “Fun” and “Useful” label self-introductions

The second independent variable, *character*, corresponded with what the online character looked like. Those in the “human” conditions interacted with a photograph of a human (the character “Ian” from buy.com), while those in the “cartoon” condition interacted with a cartoon of a person (the “The Genius” character from Microsoft). (See Figure 15.)



Figure 15. “Ian” and “The Genius” characters

4.1.2 Participants and Procedure

Students [$N = 48$] randomly selected from a large undergraduate lecture course, were randomly assigned to one of the four conditions, with gender balanced across conditions. All participants received class credit for their participation. They completed the experiment on the web, from their residences (mostly on-campus dormitories).

Each participant was given a task scenario (from Huang, Lee, Nass, Swartz & Young, 2000): “[Y]ou have just graduated from Stanford and are moving to another city. You are going to buy a lot of stuff for your new apartment rather than bringing your old stuff with you.” As in Huang *et al.*, this scenario was potentially relevant to all students, regardless of gender. Participants were told to explore a simulated e-commerce website (“abc.com”; see Figure 16), and buy at least two items for their apartment using a provided “gift certificate.” The agent, in a frame to the left, provided additional guidance and advice, although it provided no additional functionality—that is, it was possible for participants to complete the assignment without ever referring to the agent. This was intended to mirror the behavior of the Microsoft Office Assistant, which offers guidance which is mostly redundant (that is, the same help facilities can be accessed elsewhere).

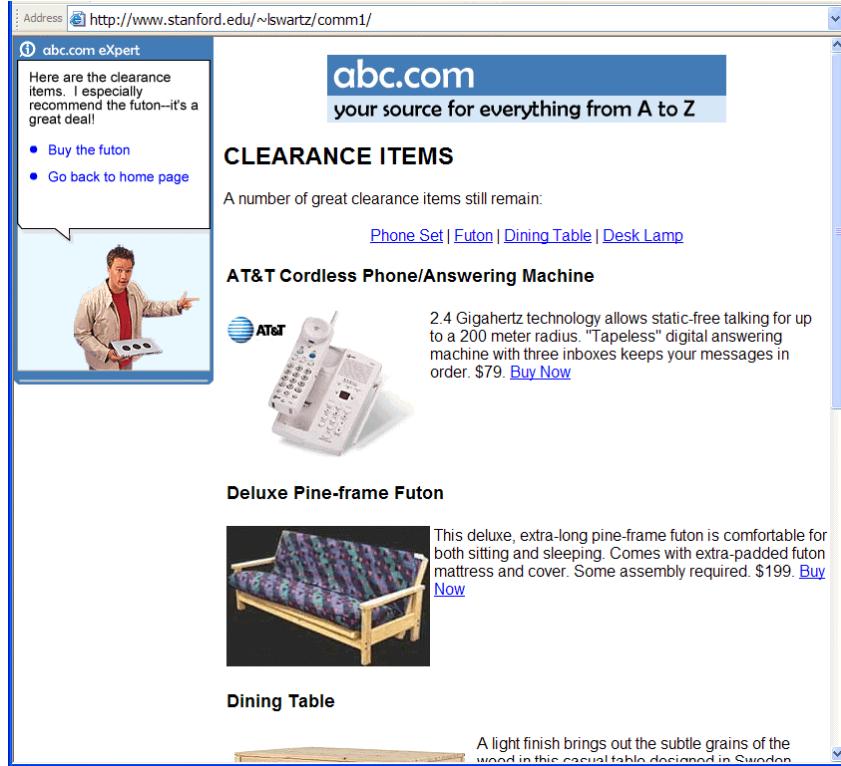


Figure 16. Sample web page from experiment

4.1.3 Dependent Measures

As a behavioral measure, we recorded which items each participant bought, as well as how long each participant spent on the website.

After completing the task, participants completed an online questionnaire with independent, ten-point Likert scale questions. The questions gauged attitudes about the website, the character, the items being sold, and participants' own feelings during the experiment. The scales ranged from "very poorly" (=1) to "very well" (=10). See Appendix A for a list of the questions asked.

The results of these questions were combined into reliable indices, using a combination of theory and factor analysis. All the indices were reliable.

Website ease of use was an index composed of two items: whether the participants found the website to be easy to navigate, and whether they found the website to be easy to use ($\alpha = 0.93$).

Feeling good was an index composed of eight adjectives describing the participants' feelings while using the website: calm, comfortable, competent, engaged, happy, in control, positive, and relaxed ($\alpha = 0.90$).

All analyses below are based on a 2x2 full-factorial ANOVA. No significant differences were found for gender, major, or year in school, so they are not reported.

4.2 Results

Ease of Use

Participants in the “fun”-label condition rated the website as being easier to use than those in the “useful”-label condition ($F(1, 44) = 4.3; p < 0.05$) (see Figure 17). This effect is being driven by the relatively high mean for the human character condition with “fun” labeling. There were no significant differences based on the appearance of the character, nor were there any significant interaction effects.

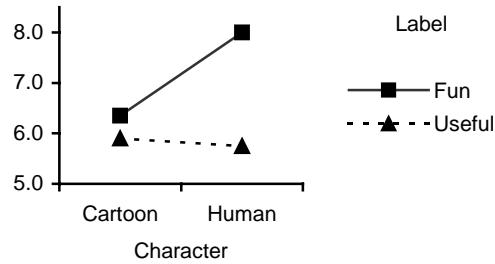


Figure 17. Website Ease of Use

Feeling Good

“Useful”-labeled participants reported feeling better during the interaction than “fun”-labeled participants ($F(1, 44) = 4.4; p < 0.05$). There was no significant effect for character appearance. There was a significant cross-over interaction ($F(1, 44) = 10.7, p < 0.01$) such that participants who saw the cartoon character labeled as “useful” reported feeling better than those who saw the cartoon character labeled as “fun”—while participants who saw the human character labeled

“fun” reported feeling better than those who saw the human character labeled “useful” (see Figure 18).

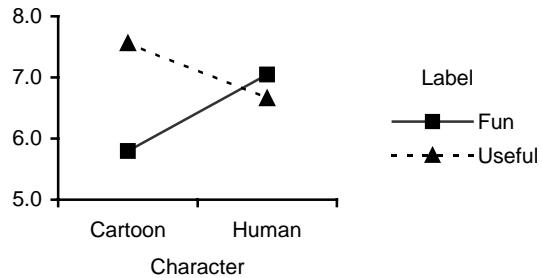


Figure 18. Feeling Good

Time Spent in Store

Finally, there was a significant cross-over interaction for the behavioral measure of how much time participants spent in the simulated e-commerce “store” ($F(1, 44) = 4.5; p < 0.05$). Participants who saw the cartoon character labeled “useful” spent more time in the store than those who saw the cartoon character labeled “fun,” while participants who saw the human character labeled “fun” spent more time in the store than those who saw the human character labeled “useful” (and Figure 19).

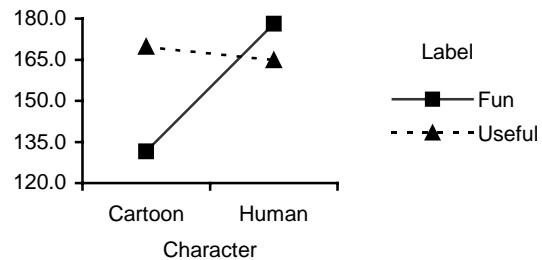


Figure 19. Time Spent in Online “Store”

4.3 Discussion

Our results confirmed that, indeed, labels do matter, even if provided by the system. Recall that the qualitative study (Chapter 3) found different perceptions based on people’s *already-existing*

cognitive labels, while in this experiment, labels were provided explicitly to participants. Most exciting is that labels not only influenced user's attitudes towards the system, but also their behavior (specifically, the amount of time they spent on the website).

That users' reactions are influenced by explicit labeling mirrors a similar study (Nass, Reeves & Leshner, 1996) in which television programming was perceived differently depending on whether it was viewed on a television labeled "Entertainment" or "News." Entertainment programming was perceived as being funnier when viewed on the "Entertainment" television, while news programming as perceived as being more informative when viewed on the "News" television.

The interaction between character and label might seem puzzling at first, as traditionally one thinks of cartoons as being more "fun" than people. However, consistency theory (Nass & Gong, 1999) presents an alternative view, in which more "person-like" characters are better perceived when they have more "person-like" characteristics (such as being "fun"). These results suggest further that the more person-like a character is, the better a "fun" label will fit that character. Consistency has been shown to be preferred in a number of agents: For example, Najmi (2002) showed that agents exhibiting consistent race (physical appearance) and ethnicity (culture, as defined by accent and greeting style) were rated higher than those with inconsistent race and ethnicity. Similarly, Isbister and Nass (2000) found that agents exhibiting consistent verbal and non-verbal personality cues were not only rated higher but influenced user behavior more.

Methodological Concerns

"The Genius" character is available in many versions of Microsoft Office; thus, people might have seen it already and associate it with the product. However, they are far less likely to associate Microsoft products with "The Genius" than with more familiar characters such as "Clippit." Since participants completed the experiment online from their residences, some factors were not controlled (introducing random error) but it also means that it has more external reliability, as the experiment was conducted in close to real-world conditions. Moreover, that

our findings were significant despite such random error would seem to highlight their importance.

It must be admitted that the two presented characters differ in more than just being a “human” or “cartoon”; the photographic human is younger and wears a tee shirt and windbreaker, while the cartoon character, which has been compared to Albert Einstein, is older and dressed in a scholarly suit. It is quite likely that the “human” character was perceived as being more consistent with the “fun” label because of its youthful appearance—particularly given the (relatively) young age of the participants, who are probably unlikely to readily associate someone of their parents’ or grandparents’ generation with “fun.” While this is, to some degree, a methodological flaw (in that more than one variable—for example, “cartooniness” and age—distinguishes the characters), note that *any* two characters will have a multitude of differences between them.

What is clear, however, is that while explicitly labeling a user interface agent can influence users’ perceptions and behavior, this decision cannot be made independent of the visual appearance of the agent. Some agents fit some labels better than others; thus, one should strive for consistency between label and appearance.

Chapter 5: Second Quantitative Study

The first quantitative study (Chapter 4) showed that explicit labeling of agents could affect users' attitudes and behaviors, especially in interaction with the kind of character presented. In keeping with the debate over anthropomorphism, what differences occur when *no* character accompanies the help agent? Also, how would changing the agent's *behavior* to be more "fun" influence the interaction?

5.1 Method

5.1.1 Design and Manipulation

The experiment was a 2 (label: "fun" or "useful") by 2 (presence of character: character or no character) by 2 (joking behavior: jokes or no jokes) balanced, between-subjects design. The first independent variable, *label*, was the same as in the first quantitative experiment, in which users were introduced to the agent as being either "fun" or "useful."

The second independent variable, *presence of character*, corresponded with whether the online agent had a character associated with it. Those in the "character" conditions interacted with a character (the same as the "human" condition in the previous experiment) while those in the "no character" condition interacted with the same text without any additional pictures. (See Figure 20.) Note that since there was no character, questionnaires referred to the agent in the no-character condition as a "box" rather than a "character."

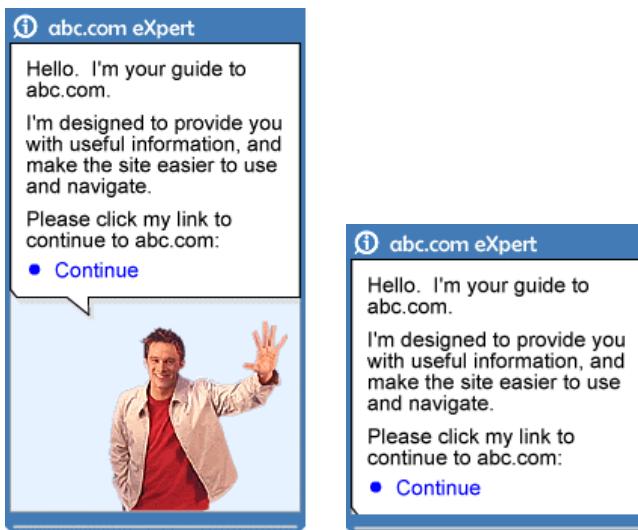


Figure 20. Character and no character conditions

The third independent variable, *joking behavior*, describes whether the character told jokes during the interaction. While the *no joke* conditions had exactly the same text as in the previous quantitative study, the *joke* conditions interspersed jokes into half of the text boxes. In keeping with Morkes, Kernal, and Nass (2000), care was taken to use “silly,” innocent humor, which neither deprecated the agent nor the participant. (See Figure 21 for an example.)

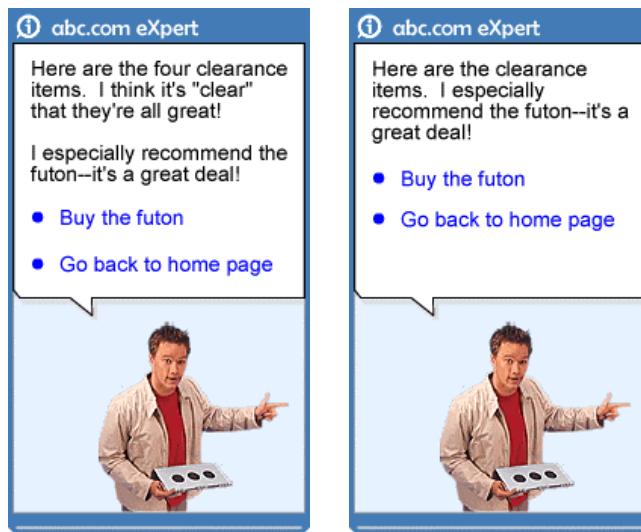


Figure 21. Joke and no joke conditions

In addition to the eight conditions in the $2 \times 2 \times 2$ matrix, ten students were placed in a ninth condition with no agent whatsoever. In this *control* condition, the space where the agent appeared in the other conditions was blank.

5.1.2 Participants and Procedure

Students [$N = 90$] randomly selected from a large undergraduate lecture course (different from the first quantitative study), were randomly assigned to one of the four conditions, with gender balanced across conditions except in three cells (in which there were six males and four females). As with the previous study, all participants received class credit for their participation and completed the experiment on the web. Other than the manipulation, the website was the same as that described in section 4.1.2.

5.1.3 Dependent Measures

The behavioral measures (items bought, time on website) and attitudinal measures (questionnaire) were the same as those for the first quantitative study (described in section 4.1.3). Those in the control condition were not asked questions about the agent, as they were not exposed to any agent, and those in the non-character condition were asked about the “box” on the left, not the “character.”

The results of these questions were combined into reliable indices—both indices used in the previous study, as well as new indices, using a combination of theory and factor analysis.

Tediumness of the agent was an index composed of two items: whether participants thought the agent was boring, and whether participants thought the agent was annoying. This index had marginal reliability ($\alpha = 0.53$).

Likeability of the agent was an index composed of ten adjectives describing the agent: engaging, enjoyable, friendly, fun, genial, intelligent, knowledgeable, personable, pleasant, and sociable ($\alpha = 0.92$).

Shop here again was an index composed of three items: whether participants agreed that the items on the site were of high quality, whether they thought items were worth the cost, and whether they would shop at a similar website ($\alpha = 0.82$).

Website reliability was an index composed of three items: whether participants found the website to be useful, reliable, and well-organized ($\alpha = 0.8$).

Website fun was an index composed of five items: whether the participants found the website to be engaging, enjoyable, fun, interesting, and likeable ($\alpha = 0.93$).

Positive rating of the website was an index composed of ten adjectives describing the website: annoying (reverse-coded), easy to navigate, easy to use, engaging, enjoyable, fun, interesting, likable, useful, and well-organized ($\alpha = 0.92$).

All analyses below are based on a 2x2x2 full-factorial ANOVA, as well as a two-tailed T-test comparing the control condition (no agent) to the other eight conditions (with agents), amplified by Dunnett's test.

5.2 Results

Number of Items Bought

“Useful”-labeled participants bought more items than “fun”-labeled participants ($F(1, 44) = 3.6$, $p < 0.1$). (See Figure 22.) This marginally significant main effect suggests that participants in the “useful” condition may have been more primed to buy things rather than to be entertained. There were no significant effects for presence of character or joking behavior.

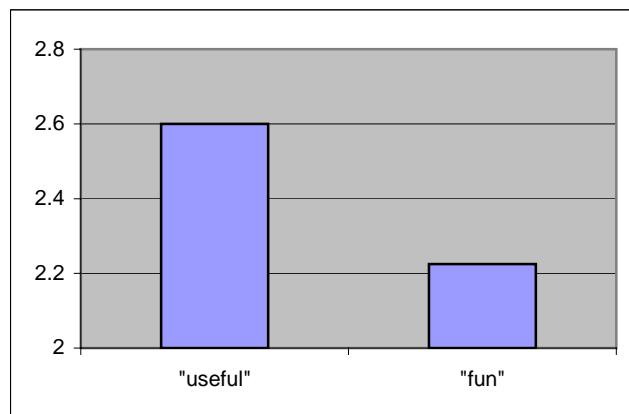


Figure 22. Number of Items Bought

Ease of Use

Participants in the various agent conditions rated the website as being easier to use than participants in the agent-less control condition ($t(88) = 1.9, p < 0.1$). (See Figure 23.) There were no significant effects for label, presence of character, or joking behavior.

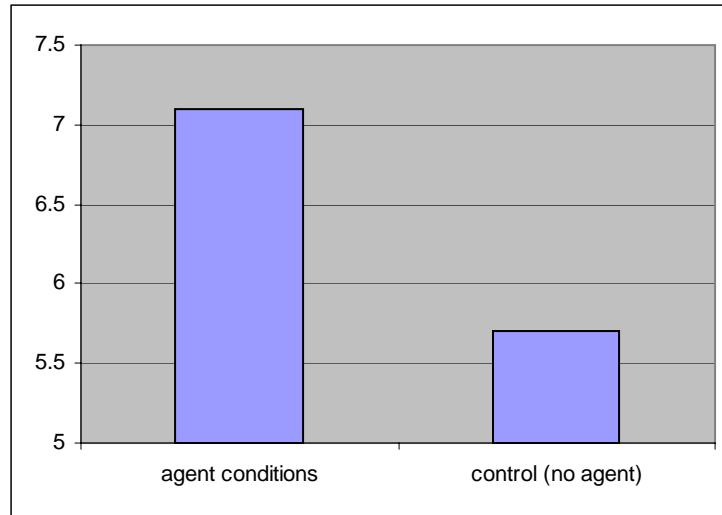


Figure 23. Website Ease of Use

Tediumness of the Agent

Participants in the various conditions without jokes rated the agent as more tedious than participants in the various conditions with jokes ($F(1, 44) = 3.0, p < 0.1$). (See Figure 24.) There were no significant effects for label or presence of character.

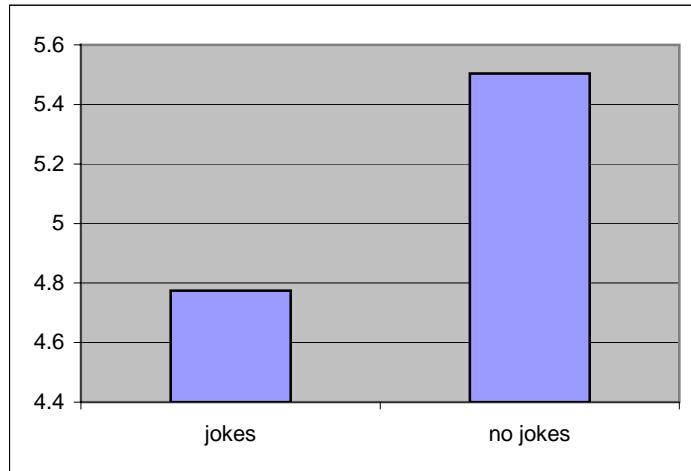


Figure 24. Tediumness of the Agent

Likeability of the Agent

Participants in the character conditions rated the agent as more likeable than those in the no character conditions ($F(1, 44) = 2.8, p < 0.1$). This effect is driven by the difference between the two conditions in which “fun”-labeled agents did not tell jokes: when such agents were *not* accompanied by a character, they were rated much lower in likeability than “fun”-labeled, non-joking agents that *were* accompanied by a character (see Figure 25).

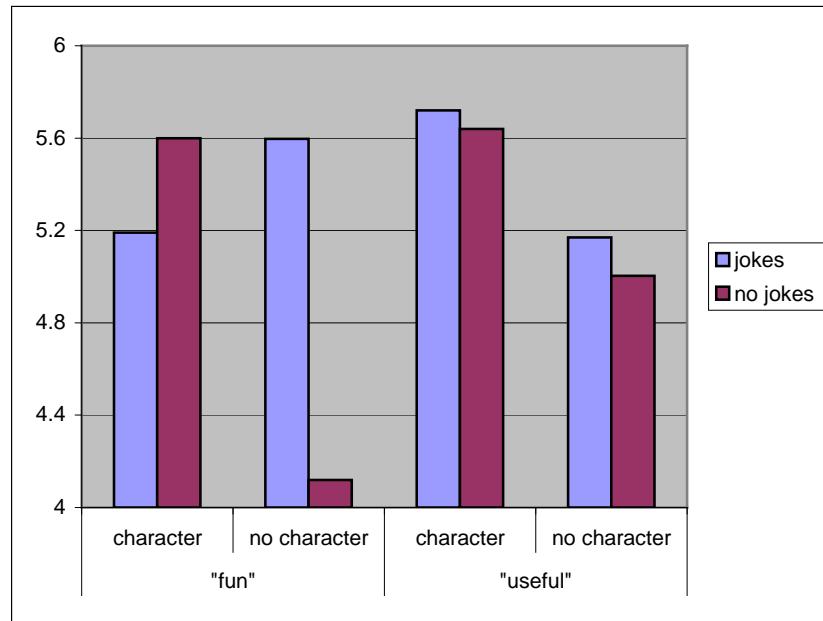


Figure 25. Likeability of the Agent

Shop Here Again

Participants in the joke conditions agreed more that they would shop at the website again than those in the no-joke conditions ($F(1, 44) = 3.6, p < 0.1$). There was also a crossover interaction with the label (Figure 26), such that participants whose agents were labeled as being “fun” and told jokes agreed more that they would shop at the website again than those whose agents were labeled as being “fun” but did not tell jokes; meanwhile, participants whose agents were labeled as “useful” agreed that they would shop at the website again about the same amount, regardless of whether the agent told jokes ($F(1, 44) = 5.6, p < 0.05$).

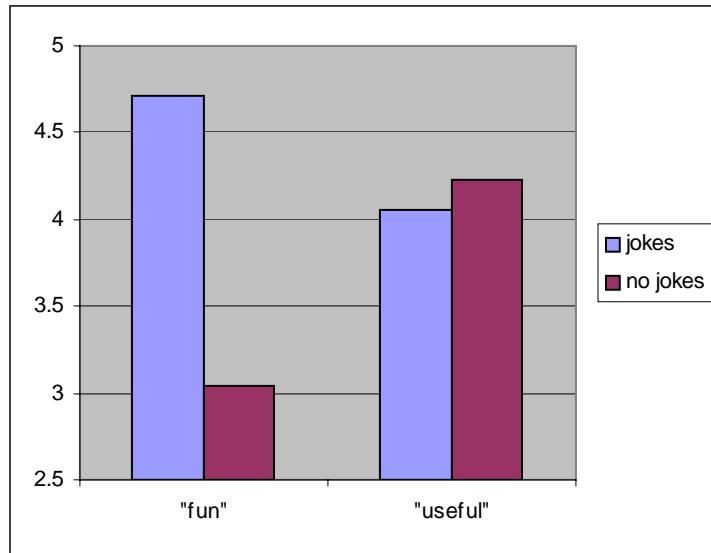


Figure 26. Users' Agreement to Shop Here Again

Website Reliability

Participants in the joke conditions rated the website as more reliable than those in the no-joke conditions ($F(1, 44) = 9.6, p < 0.005$). The participants in the control condition also rated the site as being less reliable than those in the agent conditions ($t(88) = 2.7, p < 0.01$). (See Figure 27.) There were no significant effects for label or presence of character.

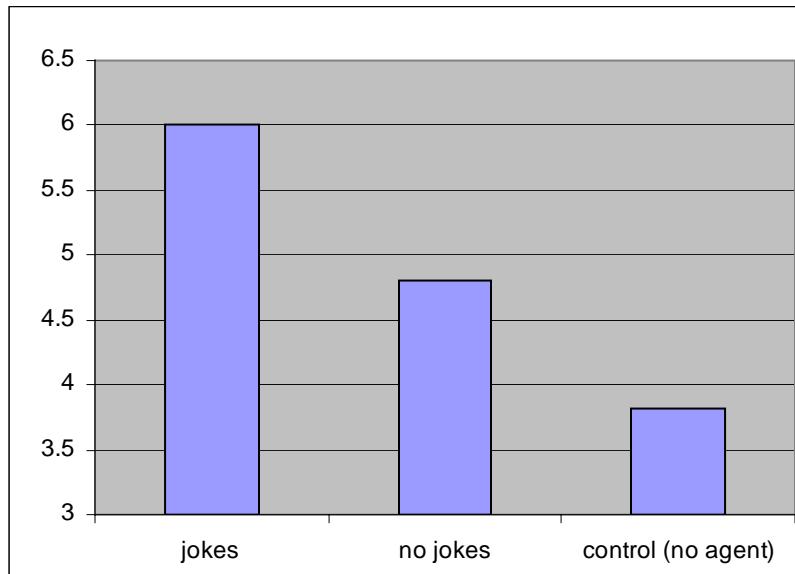


Figure 27. Website Reliability

Website Fun

Participants in the joke conditions rated the website as more fun than those in the no-joke conditions ($F(1, 44) = 7.8, p < 0.01$). There was also a significant interaction with presence of character (Figure 28), such that of participants who were not presented with a character, those who were in the joke conditions rated the website as much more fun than those in the no-joke conditions, especially compared to participants who *were* presented with a character ($F(1, 44) = 3.4, p < 0.1$). This suggests a ceiling effect, in which the presence of a character or a joking agent result in higher ratings for fun, but where the combination is not higher than either one alone.

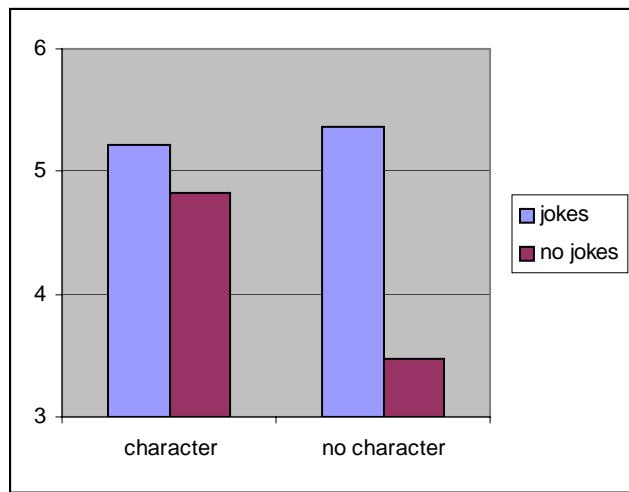


Figure 28. Website Fun

Positive Rating of Website

In a highly significant main effect, participants in the joke conditions rated the website more positively than those in the no-joke conditions ($F(1, 44) = 20.1, p < 0.00005$). There was also a main effect for presence of character, such that participants in the character conditions rated the website more positively than those in the no character conditions ($F(1, 44) = 4.9, p < 0.05$). However, this main effect is largely driven by the interaction between presence of character and joking behavior, in which, like the Website Fun index, characters and jokes each raise the positive ratings, but in combination they are no higher than either one alone ($F(1, 44) = 4.0, p <$

0.1). Finally, there was a significant three-way interaction for people's positive ratings of the website ($F(1, 44) = 5.4; p < 0.05$). While participants in the "useful" conditions rated the website more positively if the character told jokes, there was little difference based on whether a character was present. However, participants in the "fun" conditions rated the website more positively if the agent either told jokes or was accompanied by a character; participants whose agents were labeled as "fun" but neither told jokes nor had a character were rated the lowest of any condition. Participants in the agent-less control condition also rated the website less positively than those in the other conditions ($t(88) = 2.5, p < 0.05$).

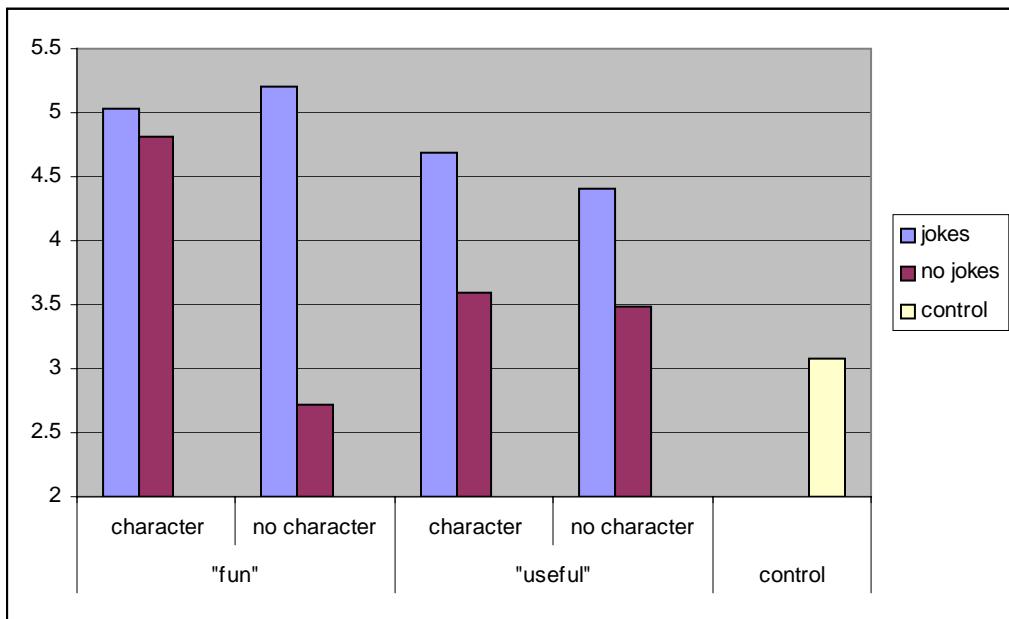


Figure 29. Positive Rating of the Website

5.3 Discussion

Once again, labels made a significant difference in people's experience of the system—although, once again, that difference depended on the kind of character presented. Users tended to buy more items when the agent was labeled as "useful," suggesting (as described above) a priming effect that users are prepared for acting rather than being entertained. However, in all other results, labeling affected users' reactions in interaction with other factors. Most strikingly, it appears that if an agent is labeled as "fun," it is important to have some fun element—either a

character with a fun appearance, or exhibiting fun behavior (such as telling jokes). In short, if one promises fun, one must deliver on that promise.

The results also confirm Morkes *et al.*'s (2000) finding that humor can increase people's attitudes towards a human-computer interaction. Agents which joked were rated as less tedious and more likeable; their websites were rated more reliable, fun, and positive, and users agreed that they would shop at the website again more. Again, this had its most profound effect in the case where there was no character, yet the agent was labeled as "fun."

The presence and absence of a character—or any agent at all—also had interesting effects. As noted above, the mere presence of a character sufficed to have a successful "fun"-labeled interaction, regardless of the agent's joking behavior. There was also a general trend for agents with characters to be perceived as more likeable than characters without, suggesting that agents with characters will be perceived more positively than those without. Finally, the control condition, without any agent, was rated as less easy to use and less reliable than the various conditions with agents. Even though the agent added no actual features to the site, it seems that, at least in this domain, the mere presence of an agent can positively influence the user's experience.

Methodological Concerns

While the no character condition did not have a character—anthropomorphic or not—one could argue that the agent's text presentations were, nevertheless, anthropomorphic: For example, the agent refers to itself as "I" and "me." Comparing the agent's current text to less personal text (for example, text that avoids first person pronouns) would make for an interesting future experiment; however, this study suggests that anthropomorphic language alone is not enough to classify an agent as "fun"—it must also have either a picture of a character, or engage in "fun" behavior (in this case, telling jokes). Interestingly, the "fun" behavior gives all the benefits that a character would, suggesting that, with the right behavior, many character-based user interface agents could be changed to simpler, less distracting text boxes. Indeed, Lee and Nass (1999) showed that text boxes were rated higher and were more effective in changing user behaviors than stick-figure characters. (Of course, this depends on the role of the agent: if the agent is

required to signal complex conversational turn-taking, it may be difficult or impossible to send these nonverbal cues to the user without resorting to a pictorial character.)

It might also be argued that the control condition was somehow odd, in that it had a blank space where the agent appeared in other conditions. While it is indeed unusual to have a large blank space on a website, replacing that blank space with other content would have added confounding variables, making the condition an ineffective control. Nevertheless, it would be interesting to run a study comparing various kinds of agent help to non-agent “extraneous” user interfaces.

Finally, it must be noted that, since users only interacted with the website once, one cannot tell what the reactions would be to the agent’s behavior and appearance over time. In the joke conditions, this agent always told the same jokes at the same times; for other often-used interfaces, it might be necessary to vary the jokes and timing to preserve the positive effects of humor.

Chapter 6: Conclusions

The Microsoft Office Assistant, both because of its unpopularity and ubiquity, makes for an interesting lens through which to look at the larger issue of user interface agents. The following are some design conclusions that would apply not only to a redesign of the Office Assistant, but to designing any user interface agent:

- Consider the agents' task in its social element (for example, beginners may want to rely on more experienced users for help and guidance—how can one facilitate this?).
- Agents should obey human rules of etiquette as much as possible (if one doesn't like a person who disobeys these rules, one will especially dislike a computer agent that disobeys them!).
- Explore ways to use the agent to teach users skills to make them more self-sufficient (thus allowing users to retain a sense of control over the program).
- Carefully introduce the agent so as to realistically showcase its best features—and be sure that the appearance and behavior are consistent with that introduction (for example, if one calls the agent “fun,” there should be something fun about it!).
- Study whether it is beneficial to use characters or agents at all (in some cases, a less anthropomorphic agent, or no agent at all, may provide the same benefits with less costs).

If one wished to draw a single lesson from this research, it might be that designing effective user interface agents is hard. Many factors—task, situation, behavior, appearance, label—influence users' responses. However, there seem to be sufficient benefits to using such agents to justify continued research to explore how these factors work. Moreover, by better understanding how we interact with agents, we may better understand how we interact with each other.

References

- Bickmore, T. (2002). When etiquette really matters: relational agents and behavior change. *Proceedings of the AAAI Fall Symposium on Etiquette for Human-Computer Work*, November 15-17, Falmouth, MA, 9-10.
- Bickmore, T. & Cassell, J. (2001). Relational agents: a model and implementation of building user trust. *CHI 2001 Conference Proceedings*, 3(1), 396-403.
- Blikstein, P. (December 2000). A new transparency: the expensive and expansive cultural dimension of user interfaces. <http://web.media.mit.edu/~paulo/courses/systems.htm>
- Brennan, S. E. (1990). Conversation as direct manipulation: an iconoclastic view. In Laurel (ed.), *The Art of Human-Computer Interface Design*, Reading, MA: Addison-Wesley, 393-404.
- Brennan, S. E. & Ohaeri, J. O. (1994). Effects of messages style on users' attributions toward agents. *Proceedings of the ACM CHI '94 Human Factors in Computing Systems: Conference Companion*, Boston, 24-28 April 1994, 281-282.
- Cole, R., Massaro, D. W., de Villiers, J., Rundle, B., Shobaki, K., Wouters, J., Cohen, M., Beskow, J., Stone, P., Connors, P., Tarachow, A., & Solcher, D. (April 1999). New tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children. *Proceedings of ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education*, London, UK.
- Cuneo, A. (May 16, 1988). Hot raisins: it's licensed products that bring big bucks. *Advertising Age*, p. 30.
- 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.
- Don, A., Brennan, S. E., Laurel, B., & Schneiderman, B. (1992). Anthropomorphism: from ELIZA to Terminator 2. *Striking a Balance: Proceedings of the 1992 ACM/SIGCHI Conference on Human Factors in Computing Systems*, New York: ACM Press, 67-70.
- Doyle, P. (May 1999). When is a communicative agent a good idea? Workshop on Communicative Agents of the *Third International Conference on Autonomous Agents*, Seattle WA.
- Dubberly, H. & Mitsch, D. (1987). *Knowledge Navigator* (film). Apple Computer. Available as part of ACM CHI 1992 Special Video Program.
- Gates, B. (1995). Speech at Lakeside High School.
<http://www.microsoft.com/billgates/speeches/education/billglake.asp>

- Heckerman, D. & Horvitz, E. (July 1998). Inferring informational goals from free-text queries: a Bayesian approach. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI, pp 230-237. Morgan Kaufmann: San Francisco.
<http://research.microsoft.com/~horvitz/aw.htm>
- Horvitz, E. (1999). Principles of mixed-initiative user interfaces. *CHI 99*, 15-20 May 1999, 159-166.
- Horvitz, E., Breese, J., Heckerman, D., Hovel, D. & Rommelse, K. (July 1998). The Lumière project: Bayesian user modeling for inferring the goals and needs of software users. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI, pp 256-265. Morgan Kaufmann: San Francisco.
<http://research.microsoft.com/~horvitz/lumiere.htm>
- Huang, A., Lee, F., Nass, C., Paik, Y. & Swartz, L. (2000). Can voice user interfaces say "I"? An experiment with recorded speech and TTS. Unpublished manuscript.
<http://www.stanford.edu/~nass/comm369/pdf/IMeVs.Impersonal.pdf>
- Isbister, K. & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53(1), 251-267.
- Flannery, M. & Merrill, D. (2000). *Faces and voices: social and cognitive effects of new interface paradigms*. Honors Thesis, Symbolic Systems Program, Stanford University.
- Fogg, B. J. & Nass, C. (1997). Do users reciprocate to computers? *Proceedings of the CHI Conference*, Atlanta, GA.
- Gay, L. & Lindwurm, D. (1985). Unpublished student project, University of Maryland.
- Kiesler, S., Sproull, L., & Waters, K. (January 1996). A prisoner's dilemma experiment on cooperation with people and human-like computers. *Journal of Personality and Social Psychology*, 70 (1), 47-65.
- Lanier, Jaron. (July 1995). Agents of alienation. *Interactions*, July 1995. Also in *Journal of Consciousness Studies*, 2(1), 1995, 76-81.
- . (November 1996). My problem with agents. *Wired*, 4(11).
<http://www.wired.com/wired/archive/4.11/myprob.html>
- Laurel, B. (1991). *Computers As Theatre*. Reading, MA: Addison-Wesley.
- . (1990). Interface agents: metaphors with character. In Laurel (ed.), *The Art of Human-Computer Interface Design*, Reading, MA: Addison-Wesley, 355-365.

- Lee, E.-J. & Nass, C. (1999). Effects of the form of representation and number of computer agents on conformity. *Proceedings of the CHI 99 Conference*, Pittsburgh, PA, 238-239.
- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: affective impact of animated pedagogical agents. In S. Pemberton, Ed. *Human Factors in Computing Systems: CHI'97 Conference Proceedings*, 359-366.
- Lester, J. C., Stone, B. A., Converse, S. A., Kahler, S. E., & Barlow, S. T. (1997). Animated pedagogical agents and problem-solving effectiveness: a large-scale empirical investigation. In B. Du Voulay & Mizoguchi, Eds. *Proceedings of the 8th World Conference on Artificial Intelligence in Education*, 23-30.
- Markoff, J. (July 17, 2000). Microsoft sees software “agent” as way to avoid distractions. *New York Times*, Technology Section.
<http://www.nytimes.com/library/tech/00/07/biztech/articles/17lab.html>
- McBreen, H., Shade, P., Jack, M. A. & Wyard, P. J. (June 2000). Experimental assessment of the effectiveness of synthetic personae for multi-modal e-retail applications. *Proceedings of the Fourth International Conference on Autonomous Agents*, 39-45.
- Miller, C. A. & Funk, H. B. (2001). Associates with etiquette: meta-communication to make human-automation interaction more natural, productive and polite. *Proceedings of the 8th European Conference on Cognitive Science Approaches to Process Control*. September 24-26, 2001; Munich.
- Moon, Y. (1998). When the computer is the “salesperson”: Computer responses to computer “personalities” in interactive marketing situations. Working Paper No. 99-041, Harvard Business School. Boston, MA.
- Moon, Y. & Nass, C. (1996). How “real” are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication Research*, 23(6), 651-674.
- . (1998). Are computers scapegoats? Attributions of responsibility in human-computer interaction. *International Journal of Human-Computer Studies*, 49(1), 79-94.
- Morkes, J., Kernal, H. & Nass, C. (2000). Effects of humor in task-oriented human-computer interaction and computer-mediated communication: a direct test of social responses to communication technology theory. *Human-Computer Interaction*, 14(4), 395-435.
- Moundridou, M. & Virvou, M. (2002): Evaluating the persona effect of an interface agent in a Tutoring System. *Journal of Computer Assisted Learning*. 18(3). 253-261.

- Najmi, S. (2002). *Social identity, race, and ethnicity in computer-based agents: implications for human-computer and human-human interaction*. Honors Thesis, Symbolic Systems Program, Stanford University.
- Nardi, B. A. & Miller, J. R. (October 1990). An ethnographic study of distributed problem solving in spreadsheet development. *CSCW 90 Proceedings*, 197-208.
- Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, 45(6), 669-678.
- Nass, C. & Gong, L. (1999). Maximized modality or constrained consistency? *Proceedings of the AVSP 99 Conference*, Santa Cruz, CA.
- Nass, C. & Lee, K. M. (2000). Does computer-generated speech manifest personality? An experimental test of similarity-attraction. *Proceedings of the CHI 2000 Conference*, The Hague, The Netherlands (April 1-6, 2000).
- Nass, C. & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.
- Nass, C., Moon, Y., & Carney, P. (1999). Are respondents polite to computers? Social desirability and direct responses to computers. *Journal of Applied Social Psychology*, 29(5), 1093-1110.
- Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, 43, 223-239.
- Nass, C., Moon, Y., & Green, N. (1997). Are computers gender-neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology*, 27(10), 864-876.
- Nass, C. & Reeves, B. (1996). *Office Assistant Introduction: Briefing on Social Interface Research*. From the personal library of Prof. Clifford Nass, Stanford University.
- Nass, C., Reeves, B. & Leshner, G. (1996). Technology and roles: A tale of two TVs. *Journal of Communication*, 46(2), 121-128.
- Nass, C. & Steuer, J. S. (1993). Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research*, 19(4), 504-527.
- Nass, C., Steuer, J. S., & Tauber, E. (1994). Computers are social actors. *Proceeding of the CHI Conference*, 72-77. Boston, MA.
- Nass, C., Steuer, J. S., Henriksen, L., & Dryer, D. C. (1994). Machines and social attributions: Performance assessments of computers subsequent to “self-” or “other-” evaluations. *International Journal of Human-Computer Studies*, 40, 543-559.

- Neff, S. (2002). *How to get whacked on the side of the head: improvisation, design and creativity*. Honors Thesis, Symbolic Systems Program, Stanford University.
- Quintanar, L. R., Crowell, C. R., & Pryor, J. B. (1982). Human-computer interaction: A preliminary social psychological analysis. *Behavior Research Methods and Instrumentation*, 14(2), 210-220.
- Reeves, B. & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Resnick, P.V. & Lammers, H. B. (December 1985). The influence of self-esteem on cognitive response to machine-like versus human-like computer feedback. *Journal of Social Psychology*, 125, 761-769.
- Rickenberg, R. & Reeves, B. (2000). The Effects of Animated Characters on Anxiety, Task Performance, and Evaluations of User Interfaces. *Proceedings of CHI 2000 - Conference on Human Factors in Computing Systems*. New York, NY, 49-56.
- Noteboom, N. (29 September 1998). Die Clippy, Die. *ZDNet AnchorDesk*.
http://www.zdnet.com/anchordesk/story/story_2589.html
- Russell, S. & Norvig, P. (2003). *Artificial intelligence: a modern approach*. Upper Saddle River, NJ: Pearson Education.
- Shneiderman, B. (1980). *Software Psychology: Human Factors in Computer and Information Systems*. Cabridge, MA: Winthrop Publishers, Inc.
- . (August 1983). Direct manipulation: A step beyond programming languages. *IEEE Computer*, 16(8), 57-69.
- . (April 1989). A nonanthropomorphic style guide: overcoming the humpty dumpty syndrome. *The Computing Teacher*, 16(7), 5. Also *Sparks of Innovation in Human-Computer Interaction*, Shneiderman, B., Ed., Ablex (June 1993), 331-335.
- . (January 1995). Looking for the bright side of user interface agents. *Interactions*, ACM Press, 13-15.
- . (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Third Edition. Reading, MA: Addison Wesley Longman.
- Shneiderman, B. and Maes, P. (1997). Direct manipulation vs interface agents, excerpts from debates at IUI 97 and CHI 97. *Interactions*, ACM, November + December, 4(6), 42-61.
- Sproull, L., Subramani, M., Kiesler, S., Walker, J. H. & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, 11, 97-124.

Steuer, J. (1994). *Vividness and source of evaluation as determinants of social responses toward mediated representations of agency*. Doctoral Dissertation, Department of Communication, Stanford University.
<http://cyborganic.com/People/jonathan/Academia/Dissertation>

Watt, S. (1998) Psychological agents and the new Web media. In Eisenstadt, M., and Vincent, T. (eds), *The Knowledge Web: Learning and Collaborating on the Net*, London: Kogan Page.

Wexelblat, A. (1998). Don't make that face: a report on anthropomorphizing an interface. *AAAI Spring Symposium Technical Report*, 173-179.

Winograd, T. (1996). Profile: Microsoft Bob. In Winograd, T. (Ed.) *Bringing Design to Software*. New York: ACM Press.

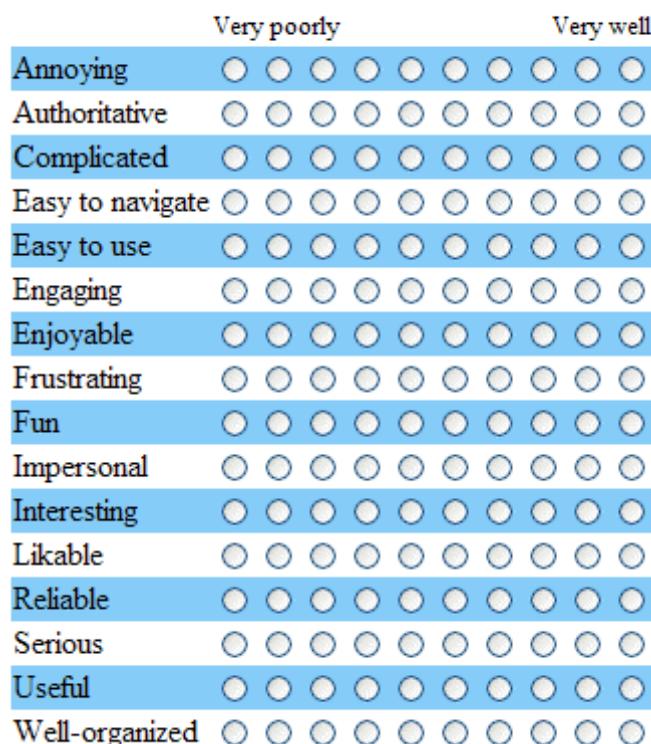
Appendix A: Quantitative Questionnaire

The following is the questionnaire used for the quantitative experiment (see Chapter 4):

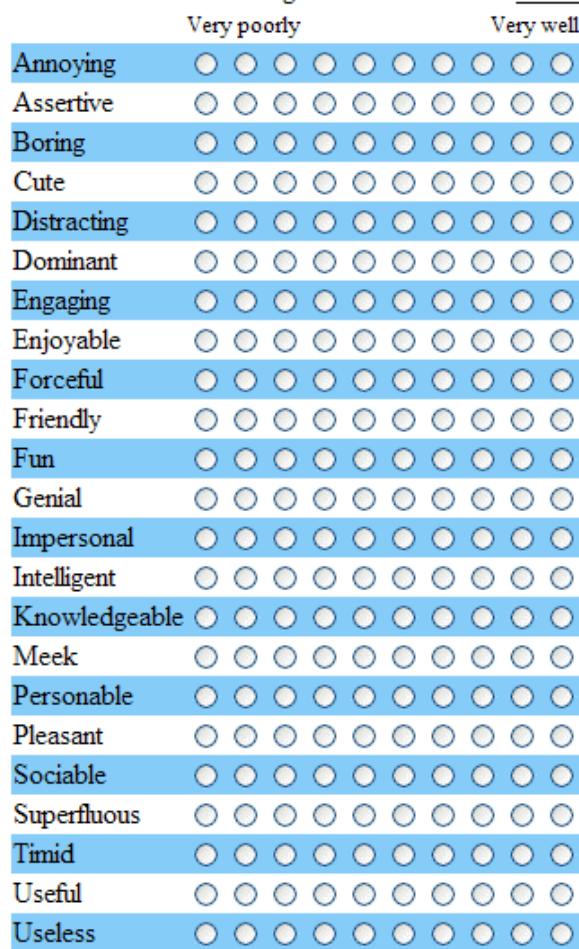
Questionnaire

Please answer all of the following questions.

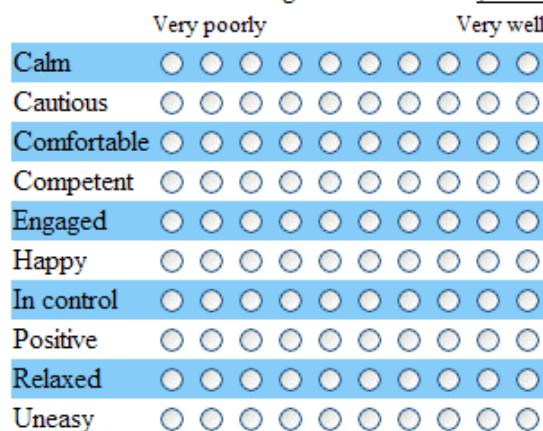
How well do the following words describe the website you just used?



How well do the following words describe the character that appeared on the left side of the screen?



How well do the following words describe your feeling while using the website?



How much do you agree (or disagree) with the following statements?

