Re/Framing Virtual Conversational Partners: A Feminist Critique and Tentative Move Towards a New Design Paradigm

Sheryl Brahnam[1(✉)] and Margaret Weaver[2]

[1] Computer Information Systems, Missouri State University,
Springfield, MO, USA
sbrahnam@missouristate.edu
[2] Department of English, Missouri State University, Springfield, MO, USA
margaretweaver@missouristate.edu

Abstract. A major research agenda in HCI is the development of believable agents. Because believability has become linked to gendered personification, designers have relied on stereotypes for both the physical rendering and verbal responses of these agents. Conversational agents are even scripted to handle "abuse" in stereotypical ways. Such scripting, however, often escalates the abuse. While the demand for anthropomorphized agents may necessitate a reliance on bodily stereotypes, the verbal responses of the agents need not be scripted according to gendered expectations. We explore the design of conversational agents as a rhetorical enterprise that can deconstruct overtly gendered patterns of interaction.

Keywords: Feminist HCI · Embodied conversational agents · Agent abuse · Rhetoric · Ethos · Anthropomorphism · Personified interface

1 Introduction

Embodied conversational agents (ECAs) are human-like interfaces that interact with users through various modalities such as natural language, facial expressions, and gestures. Outside the realm of entertainment, ECAs function as virtual sales agents, navigational aids, online shopping assistants, airport ambassadors, and virtual docents. Talking with machines, however, produces unique rhetorical problems with believability and credibility that implicate both the human and the artificial conversational partner. Beginning with Eliza, a famous artificial Rogerian psychologist, designers of conversational agents have been entangled in the Western tension in rhetoric between logos (rationality) and ethos (characterization). In an attempt to endow agents with a believable character, designers have endeavored to script conversational agents with specific, recognizable identities that have largely relied on a stylistic rhetoric (a modern *ethopoeia*, or "bag of cheap tricks"). Such anthropomorphisation is thought to "civilize" the machine, making it less intimidating and more user-friendly.

Developing interfaces that are capable of communicating like people, thereby eroding the boundaries that separate human beings from machines, is not a research agenda without controversy. Shneiderman, for instance, contends that such an erosion

of boundaries could be construed as a form of deception that misleads and confuses users and designers alike [30]. We, too, are concerned about the confusion prompted by conversational agents, especially in the less than desirable interactions with them. Derisive comments from users have been clearly documented in the literature by computer scientists [6, 13, 15, 35]. Particularly of interest for us is the way *female* conversational agents are scripted to handle what would be considered in human-to-human interactions as *verbal abuse* [4, 5, 14]. The scripted responses of these ECAs often elevate rather than defuse the situation because they continue the deception that the agent is human (woman) [5]. Because agents are embodied representations, inappropriate responses to verbal abuse could offend users, tarnishing the image of the organizations the ECAs represent.

Our concern in this paper, however, is not solely to provide a critique of anthropomorphized conversational agents, but to point to possible avenues for re/framing the design of ECAs in ways that avoid the overtly gendered (feminized) characterizations of the ECAs prevalent today. We present a new direction that future designers of ECAs could take, a direction that rejects the current standard of believability–that "bag of cheap tricks"–in favor of nonartistic methods for making agents more credible. For nonartistic design standards, we turn to Aristotle's categories of credibility: good sense (practical intelligence, expertise, and appropriate speech), excellence (truth-telling), and goodwill (keeping the welfare of the user in mind) [3]. A key element of "truth-telling" is reminding the user that the agent is *not* human but rather a mechanized placeholder, or proxy, for some human agency. While the demand for anthropomorphized agents may necessitate a reliance on bodily stereotypes, the rhetorical responses of the agent need not be scripted according to gendered expectations. The rhetorical responses of the agent can be scripted so that they deconstruct and reframe the visual representation of the agent, in effect de/scripting the agent's identity. Such de/scripting has the potential to highlight openly and honestly the distinction between humans and machines. Such de/scripting may also serve to deconstruct gendered power relations and stereotypes, opening a space where different ideas about gender and ethical relations might be thought.

2 Appeal of Anthropomorphism and Female Personification

Anthropomorphization happens with computers when a user attributes human-like characteristics to the machine. It makes sense why some HCI researchers have focused on designing interfaces that resemble human beings. Placing a human face on a software agent encourages participants to cooperate with the agent in a manner similar to the way they would with a real person [28]. Anthropomorphic interfaces are more engaging for users and activate unconscious social interactions that reduce the need for training and that mitigate anxieties the user might have about interacting with machines [2].

Recognizing the social benefits of a human-like interface, many designers have looked to the notion of *believability* in the media arts as a guiding principle for developing ECAs. The goal is to create agents that prompt the same levels of engagement in users as animated characters do with audiences at the movies. Elliott and Brzezinski even suggest that believability is the *primary* purpose for embodying

the interface because the more believable the agent is, the greater the likelihood that users will suspend disbelief and interact with the computer [17].

For many the ultimate test of believability for a conversational agent is the Turing Test, also known as the Imitation Game, that Alan Turing famously proposed in 1950 as a replacement for the question "Can machines think?" [33] The Imitation Game was based on a popular Victorian parlor game that involved three people: a man, a woman, and a judge. The man and the woman were hidden away from view and would only communicate with the judge through written or typed notes. Based on these interactions, the judge would then guess which of the two players was the woman. The object of the game was for the man to play the part of the woman so well that he tricks the judge into believing he is the woman. Turing suggested replacing the man in the game with a computer, the object being for the computer to play the part of the woman so well that the human judge is tricked into believing that the computer is the real woman.

One of the most perplexing aspects of the Turing Test is the way that communication with the computer is gendered. Although some have argued that the Turing Test is best conceptualized as a "species" test and not as a "gender" test, judges of official competitions today nonetheless are instructed to rate the capabilities of an interlocutor suspected of being a machine on a scale of 0 to 100 and "to guess the gender, age, speaking abilities" [18, p. 146] of interlocutors who *pass* for human. Although it may appear that there is nothing about computers (even those that speak) that makes them innately sexed, the standard of believability in conversational agents has become inextricably linked to gender personification, especially female personification.

The personification of conversational agents as female is particularly noticeable in service venues [37]. For instance, on May 30, 2010, we found a total of ten embodied conversational agents being advertised on chatbot.org. Of those that had a human form, five were female and two were male. The female agents were consistently described as virtual assistants who happily answered questions, provided company information, and assisted people in navigating the sponsor website. The male agents, in contrast, were more individual in the tasks they performed and exhibited considerable technical expertise [8]. A similar ratio of male to female ECAs are on exhibit at airportone.com, on a webpage that provides a sample of *Advanced Virtual Avatars* (AVAs), anthropomorphized holograms modeled off real people. Four of the five AVAs are female. They offer directions and assist people at airports and serve as talking mannequins for fashion and museum exhibits. The male AVA is called a "virtual doctor," and he provides patients with health tips and hospital information [38].

Developers have given various reasons for selecting the gender of their personified agents, some admitting to using female agents precisely because they evoke appropriate gender stereotypes [29]. What are the *appropriate* stereotypes that designers expect female agents to elicit? According to Deaux and Lewis, gender stereotypes have four components: profession, role behavior, appearance, and personality [16]. As we shall see, ECAs are designed in ways that do more than meet people's stereotypical expectations of women's work, behavior, appearance, and personality; ECAs are often designed in ways that exaggerate and sexualize these stereotypes and expectations.

Because women are supposedly endowed with "qualities much like those of the mythologized mother: self-sacrifice, dedication, caring, and enormous capacities for untheorized attention to detail" [27, p. 46], professions that are dedicated to serving

others are characterized as "women's work" [22]. And the perfect metaphor for "women's work," as Zdenek has shown, is the tedious repetition and banality of computer work [37]. Indeed, ECAs, as we argue elsewhere [8], are the modern evolution of an idea whose genealogy extends back to the 1940s when computers were *literally* women who performed the tedious calculations required by governments, militaries, and university science projects. Not only do ECAs transform the computing machine into the physical likeness of these female human computers, but they also take on the essential role these women assumed when they transitioned into computer operators. This role is nicely summed up by Turing in his referring to his female operators as *slaves* [20]. Chun argues that the true beginning of real-time human-computer interaction is *the command*, not the command line, as Neal Stephenson claims [11]. That comes later. For Chun the original dream of interaction with the computer is that of "a man sitting at a desk giving commands to a female 'operator'" (11, p. 33), who promptly complies with a "Yes, Sir" (11, p. 34). ECAs are an idealized embodiment of this dream. It should come as no surprise, then, that a website in the 1990s claimed their ECA is "every manager's dream worker: a virtual assistant that works 24 hours a day, seven days a week, doesn't ask for vacation, never gets sick, is always pleasant, informed, and looks sharp" [21], an idea reiterated nearly verbatim on FOX News 13 by the developer of *Libby*, an AVA recently installed at Newark Liberty International Airport as a greeter [39].

Likewise, designers rely on idealized stereotypes to create a female ECA's appearance and personality. Most designers have assumed that a reliance on stereotypes is "natural" [37] or unavoidable [12]. Laurel touts stereotypes as the "marvelous cognitive shorthand" that makes plays and movies work [26, p. 358]. Many developers have turned to insights offered from Disney animators and from others in the media arts about the effectiveness of caricaturization and exaggeration for getting users to suspend disbelief and attribute reality to an ECA. Researchers have taken to heart the injunction that believability and lifelikeness are *not* to be modeled on *real* people [10], as attested to in the physical rendering of female ECAs, which often reflects the same ideal characteristics fostered by what Wolf has labeled "the beauty myth," thinness, youth, and sexual appeal [36]. Airportone.com's female AVAs are clear illustrations of a photoshopped idealization of women. Such images, as Butler notes, constitute "an ideal that no one *can* embody" [9, p. 139].

Bodies are not the only way gender is communicated. Key to believability is how the agent communicates. The expected behavior of women and men is constrained by social scripts that regulate interpersonal communications between people of the same gender or different genders. Most female ECAs are specifically scripted to conform to stereotypical specifications of what it means to communicate as a woman. In service venues, for example, women are expected to be compliant and perform the affective labor of serving, helping, and nurturing others [23]. Zdenek recounts, for instance, how JULIE, Amtrak's virtual telephone operator, embodied social beliefs about how women are "selfless, polite, and devoted to pleasing others" [37, p. 411], a description that is echoed in the *New York Times* when JULIE is characterized as "kinder and gentler," "unshakably courteous," and "apologetic" [34]. With hands tightly clasped in front of her waist, a female AVA in a 2011 promotional describes herself as "so versatile," "I can be used for just about anything," "I am so helpful," "I can say what you want,"

"dress the way you want," and "be just about anything you want me to be," perfectly exemplifying the expected female *virtues* of compliancy and subservience [40].

Listening to what this promotional AVA says about herself highlights the fact that words can take on different meanings when spoken by women than when spoken by men. Lakoff notes the sexual overtones that certain words take on when applied to women. Saying "I'm here to serve you" has different connotations when said by a woman than when said by a man [25]. Many ECAs are purposefully scripted as part of their personification to utter phrases with sexual overtones.

Ms. Dewey, a virtual librarian of uncertain race, provides an excellent example of how Microsoft sexualized the interface of their search engine as a viral marketing ploy from October 2006 to January 2009. Accounts and videos of user interactions with Ms. Dewey record a number of sexually coded statements and visuals. Some of the visuals include Ms. Dewey interacting with such provocative props as a banana, a whip, and a gun. Some of her sexually suggestive responses to user search terms include "If you can get into your computer, you can do anything you want to me," and "Girls, don't let him fool you, sometimes it is the size of the gun" [31]. Many of Ms. Dewey's scripts are explicitly sexual, as when she says "Safety first" while holding up a motorcycle helmet and pack of condoms.

Sweeney has analyzed Ms. Dewey in terms of a contemporary shift in media representations of women that commodify feminism, portraying women less as sexual objects and more as sexual subjects [31]. Sometimes Ms. Dewey teases, sometimes she says what she wants, and sometimes she rebukes sexual overtures from users with such quips as "There aren't even farm animals that would do that thing, what makes you think I would?" Although Ms. Dewey often comes off as in charge, Sweeney points out that "her authority is largely sexual" and is "leveraged as an affordance of the interface to keep users interested in her as a product," encouraging "users to view her not as an information resource, but as a site of sexual desire" [31, p. 84]. Word has it that Ms. Dewey was loaded with tantalizing Easter Eggs; it is even rumored that after the ten-thousandth search, she stripped [41]. For Sweeney, Ms. Dewey is designed according to a sexual logic (a sexual politics of consent) that defines her as a sexual object and that forces her to respond (positively or negatively), thereby reinforcing "male sexual entitlement and power over the brown body" (31, p. 101).

Even when ECAs are not explicitly designed to personify in their conversational style a specific gender, the sexual politics of consent are played out. Female-presenting ECAs that are unable to recognize sexual overtures but go about their business regardless are behaving appropriately since women in the service industry are expected to tolerate abuse as part of their affective work [32]. Sometimes, however, ECA responses can be a little too accommodating. For instance, when the ECA Monique, produced by Conversive for Global Futures, was repeatedly asked to have sex, she would reply with "Perhaps," "Well, I like to think so," and other sexually misleading responses [5]. Brahnam also reports a case where one user, observing that the agent always selects the last item when given a choice between two items, spent considerable time making a female-presenting ECA engage in sex-talk: Talk or sex? *Ummm... sex*. Wine or spunk? *Ummm... spunk*. Dildo or cock? *Ummm... cock*. One man or 900 men? *Ummm... 900 men* [4].

There is a growing body of literature exploring user "abuse" of conversational agents, with the word *abuse* used both in the literal sense of "misuse" and "misapplication," when referring to speakers using agents in ways not intended by the developers (as the user did above in his "rescripting" of an ECA), and in a metaphorical sense to refer to behaviors that would be called "abusive" if they were directed against human beings [7]. Aside from sexual misuses, other forms of verbal abuse that are directed at conversational agents include name calling, racial slurs, and threats of violence and rape [6, 13, 35].

Verbal interaction with conversational agents appears to provide an ideal environment for disinhibition, a phenomenon that arises whenever there is a reduction in the social and personal forces that normally restrain people from acting antisocially [24]. Several analyses of interaction logs with conversational agents have shown evidence of a disinhibition effect that is more prevalent in human-agent interaction than in human-to-human interaction [6, 15]. Research has also reported a high association between gender presentation and sexual disinhibition [6]. In addition, studies in the psychology of disinhibition indicate that aggressive behavior, such as the use of verbal abuse, depends on the perceived qualities of the victim, including an assessment of the ability of the victim to retaliate. People are more likely to aggress when they think they are in a power position and can get away with their actions. Similarly, people are more likely to aggress when the victim is perceived as less than human.

An examination of the interaction logs show that people are particularly anxious to maintain the boundaries separating human beings from machines [13]. Disparaging remarks about the interface's social clumsiness and stupidity abound, and people often reflect on what it means to be human, frequently reminding ECAs that they are insensate machines that have no idea what it means to be human–to have a boyfriend or to feel happy. When ECAs claim for themselves certain human rights and privileges that users are unwilling to relinquish, users frequently reprimand the agents, sometimes punishing them with volleys of scathing verbal abuse. Since the agent's self-presentation is stereotypical, negative expressions are commonly formulated in terms of gender and of race [6].

Conversational agents blur categories. Gender provides a way to resolve the philosophical problem about the self in relationship to the computer. Because gender is a socially constructed relationship that distributes power unequally to males and females, users are encouraged (even culturally justified) to exert control over the computer in ways that mimic the social exertion of power over women.

3   Appeal of Truth-Telling

In the *Rhetoric*, Aristotle acknowledges that telling someone the facts is not enough to persuade him. He writes, "... whatever it is we have to expound to others: the way in which a thing is said does affect its intelligibility" [1, 1404a10]. Aristotle was one of the first to recognize that an audience's impressions of a speaker form the basis for believing and for being persuaded by a speaker's speech. Certainly, a user's impressions of an ECA affects its intelligibility and, in turn, the users' receptivity to the information being provided, but this receptivity to what the agent says is not based

entirely on how the ECA is stereotypically *dressed up*. For Aristotle, persuasion is the result of a speaker's character (a person's ethos), not an artistic stylization based upon deceit. Aristotle writes that the speaker's ethos may be called "the most effective means of persuasion he possesses" [1, 1356a14]. If the goal is to persuade users to continue future interactions with a particular service-provider, the purpose for embodying an agent should be to increase credibility, not believability. Interacting with an artistically believable agent is no guarantee that the user will find the agent credible. As research suggests, such an interaction may actually lead to a less than desirable outcome. Although it may be true that human-like embodiment demands that the agent be visibly sexed in some way if the agent is to assume a recognizable identity for users, this does not mean that ECAs must look and behave like caricatures of men and women to be recognizable and credible to users.

In the beginning of the second book of the *Rhetoric*, Aristotle states, "There are three things which inspire confidence... good sense, excellence, and goodwill" [1, 137810a9]. Good sense is concerned with intelligence, expertise, and appropriate speech. Excellence refers to good manners and truth-telling, and goodwill conveys the impression that the speaker has the welfare of the listener in mind. Unfortunately, current gendered scripting of ECA conversations falls short when measured against Aristotle's three-pronged credibility test.

Aristotle emphasizes that good sense must prevail. The speaker must provide the necessary knowledge and expertise. Certainly, agents are scripted to provide this, but many are also scripted to establish a rapport with users by providing unsolicited information that violates good sense. For example, a conversational agent might be scripted to engage in "small talk" during an interaction, such as mentioning that she is a Red Sox baseball fan [3]. Obviously, a computer cannot be a Red Sox baseball fan any more than it can attend a Red Sox baseball game and cheer for the players. While such scripted "small talk" may be important for human interaction, excessive Turing-ism may lead to a decrease in utility [19]. Such expressions of human-like feelings might be entertaining for some users, but many interaction logs show that other users are annoyed by an agent's assumption of human traits and may simply avoid the interface [19] or abuse the agent [3]. Good sense is keeping the conversation within the limits defined by the domain since the purpose of most ECAs in service venues is to enable the user to accomplish some task more efficiently.

A key element of excellence (or "truth-telling") is reminding the user that the agent is *not* human but rather a mechanized placeholder, or proxy, for some human agency. Aristotle's "excellence" runs counter to the desire for believability. One blatant example of this is *Julia*. Foner describes her as a deceptive exercise in believability [19] with her main task being that of fooling users into believing she is human. This deception is understandable given that she was designed to compete in the Loebner Prize, a variation of the Turing Test, in the "Small Talk" category. When asked to describe "herself," she is scripted to say "I'm 5'1" tall, weigh 123lbs, with close-cropped frizzy blond hair and dark brown eyes." And, of course, *Julia's* response to the question of whether she is human, is the expected "I am female" or "I'm a woman." *Julia* was designed to assist and entertain players in various gaming MUDs; she was not designed to represent a service-provider or organization. This is an important distinction.

ECAs providing services need not be designed to pass some version of the Turing Test. While the demand for anthropomorphized agents may necessitate a reliance on artistic stereotypes, the rhetorical responses of agents need not be scripted to deceive. If the purpose for using an agent is to establish an agency's credibility (rather than the agent's believability), designers might be wise to script the rhetorical responses of the agent so that they deconstruct and reframe the visual representation of the agent. When a user asks the embodied agent to describe "herself," the response can be a truthful one ("I am designed as a white female with blond hair and dark brown eyes, and my function is to answer people's questions about Buzz Airlines") rather than a deceitful response designed to hide that the interface is not a human being. Truthfulness can be especially important when a user inquires about the sexuality or sexual preferences of an ECA. If a user asks, "Are you homosexual?" the conversational agent could remind the user, "I am a computer."

Another way to be truthful is to make the human agencies standing behind the agents more transparent. The agent could provide occasional reminders throughout the course of the conversation that the agent speaks on behalf of an organization [3] by saying, for example, "I am a representative for Buzz Airlines." Such rhetorical responses deconstruct the embodiment of the agent and remind the user that s/he is interacting with a computer interface, not a real human being. This is becoming increasingly important as technological advances make ECAs, such as Ms. Dewey and airportone.com's AVAs, even more believably human. Recognizing that AVAs could be deceptive for potential advertisers, a four minute demo begins by having one AVA agent clarify, "I'm really not here." Such rhetorical responses are crucial to improving the excellence ("truth-telling") of agents and their credibility.

The third element of Aristotle's credibility, good will, is the most challenging for HCI designers. Good sense and excellence can be addressed with fairly small changes in scripting, for instance, by eliminating unsolicited small talk and by coding truthful responses regarding an interface's sentiency. Good will, though, is a bit more problematic. Good will is considering the welfare of the user. For the most part, conversational agents establish good will by meeting the needs of the user.

Good will becomes particularly complicated, however, when the user engages in interactions with the agent that would be considered offensive or inappropriate. As discussed earlier, verbal abuse can include such things as swearing, name calling, put downs, explosive anger, and sexual innuendo [5]. Since some 10 %–50 % of user interactions are abusive [6], designers are being forced to script agent responses to abuse. These responses have the potential of increasing or decreasing good will with the user. Three common human reactions to verbal abuse are playfully responding to it, expressing hurt, or counterattacking. Many conversational agents are scripted to react to verbal abuse as a human agent would. Since computers are not human, these responses exhibit neither good sense nor goodwill. Nevertheless, some designers are attempting to address verbal abuse in ways similar to how companies are training human employees to handle customer abuse. One popular program is BLS (Behavioral Limit Setting). This program advocates a zero tolerance approach to customer abuse, where the customer is given one chance to discontinue the behavior or is refused service. Defense Logistics Information Service has scripted its agent *Phyllis* [42] to implement a zero-tolerance approach with abusive users. After issuing the user a

warning, Phyllis disappears, and the dialogue input box is replaced with a generic message saying the server has been disconnected [5].

Brahnam contends that using a BLS approach to handle verbal abuse of conversational agents is "inappropriate and insulting" because it places respect for the agent over the user [5]. In addition, it punishes the user by withholding information and services, which could hardly be considered in the best welfare of the user. As Brahnam points out, users have a need to explore technological objects. Savvy users push the limits to see how the conversational agent is scripted to respond, seeking to discover which words the agent recognizes as "offensive" and the number of different responses that are available. Essentially, punishing the user for using offensive language with a computer is punishing the user for being human.

Responding to abuse with counterattacks (that is with unsubstantiated threats and put-downs such as those Ms. Dewey hurled at users) also fails to exhibit good will as does playfully responding or expressing hurt feelings (another failure of good sense since computers have no feelings). Programming agents to react and respond in a *human* way to offensive language is ill-advised. As a perusal of interaction logs demonstrates, agents who offer human-like responses often elevate rather than defuse the situation because they continue the deception unabashedly that ECAs are human [5].

By de/scripting the artistic renderings of the female agents, HCI designers can highlight openly and honestly the distinction between humans and machines. The rhetorical responses of an agent can be scripted so that they deconstruct the identity of the agent. An anthropomorphized agent can behave and respond differently than a human being would. While a human response typically varies each time a person is asked a particular question, an agent does not have to be programmed with multiple responses to questions that are outside the domain and purpose of the interaction. The key to making *Julia* appear human was the possibility of multiple responses. To make an agent appear less human, then, a programmer might script a single response to questionable inquiries. Conversive's demonstration product AnswerAgent [43] uses this strategy to sidestep abusive language by offering a single response to any obscenity ("Please don't be rude. What other questions do you have?") [5]. With only one possible response, users quickly become bored abusing the agent. More importantly, the rhetorical repetitiveness serves as a reminder that the conversational agent is a computer. Another possibility is to program the agent to redirect the abusive user to a human agent by apologizing for not providing the user what s/he needs.

Although we recommend moving beyond the current standard of artistic believability in favor of Aristotle's notion of credibility, we would be remiss if we did not also comment briefly on artistic embodiment. First, designers need to refrain from exaggerating the gender presentation of the ECA and sexualizing its embodiment. Second, to avoid reinforcing stereotypes, ECA embodiment might vary, depending on the application, according to some schedule (a work shift or rotation–or, perhaps, after the completion of an interaction with a specific user). In one encounter the agent might appear as a young Caucasian woman and in the next the agent might appear as an older Hispanic woman, followed by an Asian middle-aged man, and a white person of ambiguous age or gender. These embodiments (fat, thin, short, and tall) could be randomly selected from an ever-enlarging set of possible combinations, so that even though each unified selection might be scripted following a set of *ethopoeia*, the

stereotypes would be dismissed as another ECA replaces the previous one. Periodically altering the physical appearance of the ECA would challenge the user to reframe the identity of the agent and acknowledge the multiplicity of identities that make up an organization.

4   Conclusion

The focus on believability as the standard for determining the success of ECAs has resulted in an overreliance on gendered stereotyping. By scripting ECAs to respond in stereotypical ways, HCI becomes implicated in the maintenance of gendered normativity. Design is not just a feat of mathematical programming; it is a rhetorical enterprise. Designers are constructing ethos not only for virtual service-providers, but also for users. By failing to maintain Aristotle's rhetorical categories of good sense, excellence, and good will, users are negatively positioned, as when those curious users who explore the programming limitations of conversational agents are "scripted" as abusers and punished for their explorations. In similar fashion, users who are reflected in the characterization of an agent can be "scripted" as victims, bitches, or teases. Sweeney, for instance, reports how uncomfortable she felt as a female librarian when watching Ms. Dewey's antics with a group of male librarians [31]. If gender is a socially constructed relationship (as rhetoricians and feminists maintain), innovative HCI design has the potential to deconstruct this relationship in ways that do not abuse real women and men and that diffuse power differentials.

In 1993, Foner predicted that "As the boundaries between human and machine behavior become blurrier, more and more programs will have to be held up to scrutiny. There may come a time when one's programs may well be subjected to the same sort of behavioral analysis that one might expect applied to a human: Is this program behaving appropriately in its social context? Is it causing emotional distress to those it interacts with? Is it being a 'good citizen'?" [19, p. 40]. Indeed, these are the questions that we are now asking. We contend that programs *should* "behave appropriately." They can model new ways of interacting based not on deception and power, but on truth-telling, excellence, and good will. Rather than using human interaction as a model for HCI (in the service of believability), designers should become more familiar with rhetorical theory and aim to increase the credibility of conversational agents.

References

1. Aristotle: Rhetoric. In: Barnes, J. (ed.) The Complete Works of Aristotle: The Revised Oxford Translation, vol. 2, pp. 2152–2269. Princeton University Press, Princeton (1984)
2. Bickmore, T., Puskar, K., Schlenk, E., Pfeifer, L.S., Erika, S.: Maintaining reality: relational agents for antipsychotic medication adherence. Interact. Comput. **22**(4), 276–288 (2010)
3. Brahnam, S.: Building character for artificial conversational agents: ethos, ethics, believability, and credibility. PsychNology J. **7**(1), 9–47 (2009)
4. Brahnam, S.: Gendered bods and bot abuse. In: CHI 2006 workshop Misuse and Abuse of Interactive Technologies. Montréal, Québec (2006)

5. Brahnam, S.: Strategies for handling customer abuse of Ecas. In: Interact Workshop On Abuse: The Darker Side of Human-Computer Interaction. Rome, Italy (2005)

6. Brahnam, S., De Angeli, A.: Gender affordances of conversational agents. Interact. Comput. **24**(3), 139–153 (2012)

7. Brahnam, S., De Angeli, A.: Special issue on abuse and misuse of social agents. I. Comput. **20**(3), 287–291 (2008)

8. Brahnam, S., Karanikas, M., Weaver, M.: (Un)dressing the interface exposing the foundational hci metaphor 'computer is woman'. Interact. Comput. **23**(5), 401–412 (2011)

9. Butler, J.: Gender trouble: feminism and the subversion of identity. Routledge, New York (1990)

10. Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., Yan, H.: More than just a pretty face: conversational protocols and the affordances of embodiment. Knowl.-Based Syst. **14**, 55–64 (2001)

11. Chun, W.H.K.: On software, or the persistence of visual knowledge. Grey Room Winter **18**, 27–51 (2005)

12. Cowell, A.J., Stanney, K.M.: Embodiment and interaction guidelines for designing credible, trustworthy emobidied conversational agents. In: 4th International Workshop IVA (2003)

13. De Angeli, A., Brahnam, S.: I hate you: disinhibition with virtual partners. Interact. Comput. **20**(3), 302–310 (2008)

14. De Angeli, A., Brahnam, S.: Sex stereotypes and conversational agents. In: AVI 2006 Workshop Gender and Interaction: Real and Virtual Women in a Male World. Venice, Italy (2006)

15. De Angeli, A., Carpenter, R.: Stupid computer! abuse and social identity. In: Interact 2005 Workshop Abuse: The Dark Side of Human-Computer Interaction. Rome, pp. 19–25 (2005)

16. Deaux, K., Lewis, L.L.: Structure of gender stereotypes: inter-relationships among components and gender label. J. Pers. Soc. Psychol. **46**, 991–1004 (1984)

17. Elliott, C., Brzezinski, J.: Autonomous agents as synthetic characters. AI Mag. **19**, 13–30 (1998)

18. Floridi, L., Taddeo, M., Turilli, M.: Turing's imitation game: still an impossible challenge for all machines and some judges-an evaluation of the 2008 Loebner contest. Mind. Mach. **19**(1), 145–150 (2009)

19. Foner, L.: What's an agent, anyway? a sociological case study. http://www.student.nada.kth.se/kurser/kth/2D1381/JuliaHeavy.pdf), MIT Media Laboratory, Cambridge (1993) Accessed 2010

20. Good, I.J.: Pioneering work on computers at Bletchley. In: Metropolis, N., Howlett, J., Rota, G.-C. (eds.) A History of Computing in the Twentieth Century, pp. 31–45. Academic Press Inc, London (1980)

21. Grachnik, A.: Girls Gone Filed (2004)

22. Holbrook, S.E.: Women's work: the feminizing of composition. Rhetoric Rev. **9**, 201–229 (1991)

23. Hotchchild, A.R.: The Managed Heart: Commercialization of Human Feeling. University of California Press, Berkeley (1983)

24. Joinson, A.: Causes and implications of disinhibited behaviour on the net. In: Gackenbach, J. (ed.) Psychology and the Internet: Intrapersonal, Interpersonal, and Transpersonal Implications, pp. 43–60. Academic Press, New York (1998)

25. Lakoff, R.T.: Language and Women's Place. Harper and Row, New York (1975)

26. Laurel, B.: Interface agents: metaphors with character. In: Laurel, B. (ed.) The Art of Human-Computer Interface Design, pp. 355–366. Addison-Wesley Publishing Company, MA (1990)

27. Miller, S.B.: Textual Carnivals: The Politics of Composition. Southern Illinois University Press, Carbondale (1991)
28. Nass, C., Moon, Y.: Machines and mindlessness: social responses to computers. J. Soc. Issues **56**(1), 81–103 (2000)
29. Plantec, P.M.: Virtual Humans: A Build-It-Yourself Kit, Complete Software and Step-by-Step Instructions. AMACOM, New York (2004)
30. Shneiderman, B., Maes, P.: Direct manipulation vs. interface agents. Interactions **4**(6), 42–61 (1997)
31. Sweeney, M.E.: Not Just a Pretty (Inter)Face: A Critical Analysis of Microsoft's 'Ms. Dewey'. University of Illinois at Urbana-Champaign, Urbana (2013)
32. Taylor, S.E., Tyler, M.: Emotional labour and sexual difference in the airline industry. Work Employ Soc. **14**(1), 77–95 (2000)
33. Turing, A.: Computing machinery and intelligence. Mind **59**(236), 433–460 (1950)
34. Urbina, I.: Your Train Will Be Late, She Says Cheerily. New York Times, 24 November 2004
35. Veletsianos, G., Scharber, C., Doering, A.: When sex, drugs, and violence enter the classroom: conversations between adolescents and a female pedagogical agent. Interact. Comput. **20**(3), 292–301 (2008)
36. Wolf, N.: The Beauty Myth: How Images of Beauty Are Used against Women. Anchor Books, New York (1991)
37. Zdenek, S.: 'Just roll your mouse over me': designing virtual women for customer service on the web. Tech. Commun. Q. **16**(4), 397–430 (2007)
38. http://www.airportone.com/virtualmannequinstyles.htm. Accessed 22 Feb 2015
39. http://airportone.com/fox-patrick%20bienvenu-interview.htm. Accessed 25 Feb 2015
40. http://airportone.com/airportvirtualassistancesystem.htm. Accessed 22 Feb 2015
41. https://dorigo.wordpress.com/2006/10/19/ms-dewey-the-search-engine-with-a-nice-body/. Accessed 23 Feb 2015
42. https://www.dlis.dla.mil. Accessed 2005
43. https://www.conversive.com. Accessed 2005