

# Put Your Best Face Forward: Anthropomorphic Agents, E-Commerce Consumers, and the Law

Carey E. Heckman  
Stanford Law School  
Crown Quadrangle  
Stanford, CA 94305-8610  
(650) 725-7788  
ceh@stanford.edu

Jacob O. Wobbrock  
Stanford Computer Science Department  
PO Box 5264  
Stanford, CA 94309  
(650) 325-2066  
jrock@cs.stanford.edu

## ABSTRACT

Highly believable anthropomorphic agents endanger electronic consumers. Because of concerning tendencies in human-agent interaction arising from agents' anthropomorphic qualities, consumers may unwittingly treat agents as competent, trustworthy, living counterparts. This paper concludes that developers must focus agent design on consumer welfare, not technical virtuosity, if legal and ethical perils are to be avoided.

## 1. INTRODUCTION

Anthropomorphic agents provide electronic commerce consumers with terrific benefits yet also expose them to serious dangers. Consumers are not fully aware of anthropomorphic agent capabilities and limitations. With surprising ease, overly trusting consumers may be persuaded to interact with anthropomorphic agents in a way that endangers them. Fraud and misrepresentation become easy.

Much of the potential for consumer abuse can be traced to agents' special abilities. Proponents of anthropomorphic agents cite the benefits of social interfaces, natural and comfortable human-computer interaction, conversational communication, and user-to-agent delegation [22]. Others fear that users will waive their autonomy while partnering with these "intelligent" helpers [20]. The jury is still out on the ultimate costs and benefits of anthropomorphic interface agents. Debates over their eventual success continue [31].

The power of persuasion introduces additional legal and ethical questions. Consider an agent serving as a salesperson in an electronic commerce "store" [11]. The store's management will be legally accountable if they configure the agent to defraud or deceive consumers. But recent lawsuits against tobacco and firearms manufacturers suggest the agent's designers must act responsibly or face serious legal consequences as well.<sup>1</sup>

It is no surprise that anthropomorphic agents may be created *too* well. Highly believable characters leveraging "humanness" will

<sup>1</sup> Henley v. Philip Morris Inc., No. 995172 (Calif. Super. San Francisco Feb. 1999), *appeal docketed*, No. A086991 (Calif. Ct. App 1st Div. Dec. 6, 1999) (tobacco); Merrill v. Navegar Inc., 75 Cal. App 4th 500 (1999) (firearms)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.  
Agents 2000 Barcelona, Spain  
Copyright ACM 2000 1-58113-230-1/00/6...\$5.00

increase the risk of misleading consumers. Consequently, anthropomorphic agents must be created as well as used with care.

Agent designers hold the key to averting the risks. They must understand the agent liability landscape well enough to negotiate its pitfalls. Most importantly, they must design with a consumer-protective mindset. With each feature they consider for their agent they must analyze their motivation for that capability within the framework of consumer welfare. And because the consumer is unable to know what goes on "behind the scenes" of an agent, the designer must strive to treat the consumer as an end and never as a means [17].

## 2. MISLEADING CONSUMERS WITH ANTHROPOMORPHIC AGENTS

### 2.1 Kinder, Gentler Electronic Shopping

Anthropomorphic agents can do much to revolutionize use of electronic commerce for consumer retailing.

Despite rapid growth of consumer electronic commerce over the past several years, many consumers still keep their distance. Today's World Wide Web interfaces are often graphically flashy but frighteningly unusable and therefore continue to confound consumers [18][32]. Navigating around a consumer electronic commerce site can be a chore. Consumers often must work too hard to find what they want to buy. Companies like MySimon.com have emerged to help consumers avoid precisely this.

Merchants have long recognized the value of face-to-face interaction [10]. But electronic shopping lacks this. Simple pleasantries like natural language conversation are absent and no history is formed with store personnel. Entertainment is rarely part of the e-commerce experience, though it plays a major role as "shoppertainment" in brick-and-mortar stores. In short, the e-commerce experience is far less satisfying than physical shopping.

Many hope anthropomorphic agents will make electronic commerce more palatable by restoring the attributes of real-world stores. But will they do it safely?

For example, an agent's speech and gestures can provide for more natural communication. Slower readers enjoy a reduced challenge. Speech and gesture can also deliver an emotional message that affects the consumer's subconscious as well as conscious experience.

Online agents will keep a profile of users, past conversations, likes and dislikes, in an attempt to create a personal experience for the user. It is easy for an agent to ask a user for their birthday,

record that date, and later send the user an email prodding them to visit for a birthday surprise (e.g., a \$20 gift certificate). Personalization delivered by an agent — even to the point of mimicking friendship — will very soon characterize the user-agent experience.

Animated agents are not bounded by the limitations of the real world and therefore have the ability to delight users by achieving the impossible. The cartoon favorite Wile E. Coyote has entertained more than one generation by falling off a cliff or meeting the heavy underside of a boulder, all the while providing a good laugh for the viewer. Anthropomorphic agents will employ such impossible antics as well to entertain the consumer.

In these ways agents will develop an online shopping experience that contains many of the real-world advantages. But, as in the real world, fraud, misrepresentation, deception and undue persuasion will be a reality.

## 2.2 Intentional Misrepresentation: Fraud

Every new technology applicable to commerce — telegraph, telephone, radio, television, to name a few — helps the unscrupulous or sloppy convince the unwary to make bad deals. Anthropomorphic agents are no exception.

Fraud has the gravest implications; United States national and state laws provide both criminal and civil liability.<sup>2</sup> Penalties can be severe. United States national law punishes criminal wire fraud by forfeiture of the proceeds and a fine of up to \$1,000,000 or imprisonment for thirty years, or both, with additional prison time if the fraud specifically targeted persons over age 55. Fraud can also trigger additional prosecution under the Racketeer Influence and Corrupt Organizations Act or money laundering statutes.

It is easy to imagine an anthropomorphic agent that misleads consumers by providing slanted or faulty information. A representative definition of fraud is the use of “calculated efforts to use misrepresentations or other deceptive practices to induce the innocent or unwary to give up some tangible interest.”<sup>3</sup> Thus an anthropomorphic agent scripted to assure consumers they are guaranteed a 10% return on their investment when the return in fact depends on stock market performance could suffice. Liability would also be possible for withholding pertinent information.

The victim’s gullibility (or lack thereof) plays no role. To constitute fraud, the misrepresentations or other deceptive practices only need to be “reasonably calculated to deceive persons of ordinary prudence and competence.”<sup>4</sup> The focus is on the act of misrepresentation or deception, not the victim.

Anthropomorphic agent fraud could take various forms. While the sale of goods through deceptive promotion seems most likely,

---

<sup>2</sup> United States law is used as an illustration for purposes of this paper, but laws against consumer fraud and deception are common throughout the world. The specifics of these laws vary, of course, but rarely to the point of providing an exception to the fundamental proposition that the design or configuration (or both) of anthropomorphic agents can have negative legal consequences.

<sup>3</sup> United States v. McNeive, 536 F.2d 1245, 1248 (8th Cir. 1976)

<sup>4</sup> United States v. Pearlstein, 576 F.2d 531, 535 (3d Cir. 1978)

other frauds are possible: designing or configuring an agent to prepare false application forms or purchase agreements, or filing false insurance claims.

## 2.3 Negligent and Innocent Misrepresentation

In an effort to give consumers more complete protection and make sellers more responsible, many legislatures and courts have provided for consumer recovery of damages even absent the specific intent required for fraud. Misrepresentations may be merely negligent. Although the person making the misrepresentation did not do so as part of a “calculated effort” to deceive, he or she failed to use appropriate care to avoid the deception. For example, a person scripting an anthropomorphic agent may not have intended that certain gestures would mislead the ordinary consumer, but could be reckless in using those gestures without anticipating how they might be interpreted. Negligent misrepresentation only requires proof that the person scripting the agent *should have known* that it would have this effect.

Innocent misrepresentation is still easier to prove. All that must be shown is that the act possessed a tendency or capacity to mislead. The state of mind of the agent designer is irrelevant.

A designer who builds an agent that portrays an endorsement of a product or service could be found guilty of negligent or innocent misrepresentation under U.S. federal and state law. An animated agent depicted as a celebrity golf professional might carry and swing an identifiably branded golf club. Consumers might infer that this professional therefore endorses the brand and would advise its purchase. Enforcement action could be expected if people discover the professional never uses the product or does not recommend its use.

On the other hand, having a nondescript anthropomorphic agent dramatize the use of the product or explain the product’s advantages would not constitute an endorsement. Although the statements might be inaccurate, the agent would not be misrepresenting or deceiving.

Moreover, sales personnel, and arguably anthropomorphic agents serving in that capacity, can “sales puff” without fear of legal liability. Non-expert opinions are considered part of the norm in sales communications. The reasonable consumer is expected to understand the minimal reliance appropriate for a merchant’s self-serving praise of its own products or services.

## 2.4 The Gray Area of Responsibility

When an anthropomorphic agent is part of a scheme found to defraud or deceive, questions are certain to be asked as to whether responsibility for the bad conduct belongs to the designer of the agent or to whomever configured the agent for the specific application, or both.

It may seem obvious that the person who programmed the agent with the deceptive speech (or nonverbal communication) should be considered liable. This is the person who last designed the agent’s behavior and placed the configured agent in the path of consumers. Those pointing the finger at this person are likely to recite: “Anthropomorphic agents don’t defraud consumers; those who configure anthropomorphic agents do.”

Nevertheless, the original agent designers and programmers are culpability candidates, too. They create the agent’s essential

framework and features. They decide what, if any, safeguards to include. They are in the best position to understand the problems the anthropomorphic agent can cause and to create the agent so that it avoids those problems.

Legal liability rules rely on differing ideas about responsibility. Some look to moral obligation: Anyone failing to fulfill an accepted moral obligation should be responsible for the damage society believes to be foreseeably related to that failure. Others look to economic efficiency. Under this approach, responsibility should be given to whoever is best able to foresee and prevent the damage from occurring, or if unavoidable, to whoever is best able to spread the cost of the damage among everyone who might suffer that unavoidable damage. In the case of agents, this might be a liable corporation that develops and profits from agent technology.

The capacity of the most effective anthropomorphic agents to adapt their behavior to those they encounter and to personalize their own conduct and speech complicates the responsibility analysis [15]. Perhaps neither the designer nor the person configuring the anthropomorphic agent can anticipate with certainty what the adaptive agent will do or say. Courts might resolve this predicament by declaring legally responsible *both* the designer and the person configuring the agent.

Precedent in law for other products suggest that the more sophisticated and commercially successful anthropomorphic agents become, the more likely legislatures and courts will assign responsibility for anthropomorphic agents to the initial creators. The reasoning is that the designers and creators are best able to understand what harms anthropomorphic agents can cause, to create agents that minimize those harms, and to absorb and distribute the costs of whatever harms cannot be avoided.

### 3. THE FOUNDATIONS OF A DANGEROUS ILLUSION

These fraud and misrepresentation threats to the consumer result from the special properties of human-agent interaction. This interaction is fundamentally different than human interaction with non-agent software. The differences lie in the unique illusion-creating properties of anthropomorphic agents — animation, gesture, natural language communication, emotion, personality — coupled with the tendencies of humans to personify technology, trust specialists, and form relationships.

#### 3.1 Agents Leverage “All That is Human”

The characteristics of anthropomorphic agents touted as strengths are those that make them more “believable.” Believability is a suspension of disbelief; the character becomes competent, personable, even alive. The ultimate goal of anthropomorphic agent design is to create an illusion so impenetrable that the user would feel badly doing anything that might hurt the character (like turning off the computer [4]). Anthropomorphic agent research groups<sup>5</sup> work to create agents that give a “believable performance”, suggesting the user is like an enthralled theater

<sup>5</sup> For the Stanford University Virtual Theater group, see [www-ksl.stanford.edu/projects/cait/index.html](http://www-ksl.stanford.edu/projects/cait/index.html).

For the Carnegie Mellon University Oz Project, visit [www.cs.cmu.edu/afs/cs.cmu.edu/project/oz/web/oz.html](http://www.cs.cmu.edu/afs/cs.cmu.edu/project/oz/web/oz.html).

patron [1][12][28]. In short, the user interacts as though the agent is a social intelligent counterpart: a *person* by any measure.

Agents maintain the illusion just described by *leveraging all that is human*. The illusion is an expected one: We as humans are used to interacting with humans and are even built to do so. For example, infants exhibit preferences for faces over other shapes [34]. It is no wonder that an anthropomorphic agent capable of replicating humanlike behavior should be treated socially. It is the only paradigm we humans have to go on, and it is one we are very good at. However, the fact that anthropomorphic agents are convincing could be alarming for interactions that involve exchange of goods and money if people are easily persuaded, deceived, or compelled. Special capabilities of anthropomorphic agents contribute to the illusion of life that may exploit consumers.

##### 3.1.1 Animation

Although not every anthropomorphic agent is animated, the ability to move contributes to the illusion of life. Chuck Jones, the famous Looney Toons animator, drew a direct comparison between a character’s movements and personality when he wrote, “We are how we move; in other words, our personalities” [16]. Note that it is not important for the animated body to be realistic. Realism and believability are very different. Jones thought it is more believable to animate animals than humans; humanlike agents are too familiar to users and therefore elicit unrealistically high standards for believability [16]. For designers of anthropomorphic agents this means animating a cartoon animal instead of a realistic human actually encourages users to interact socially, treat the agent as a competent counterpart, and buy into the illusion. In fact, inanimate objects are convincing. Work on agent expressiveness shows that even non-living objects such as lamps can be believable through their motion [21][30].

Undoubtedly long ago our ancestors gained the ability to associate motion with life, as it would have been impossible to survive without it. Thus it is not surprising that an animated anthropomorphic agent stirs within us a strong response to imbue other lifelike qualities. If an animated sales agent did nothing more than deliver a paragraph of persuasive text, we might be more likely to consider the information thoughtful and personal than if we read the plain text ourselves. This, after all, is the rationale behind testimonial-style advertising. Animation is the beginning, but by no means the end, of what erects the illusion of life.

##### 3.1.2 Gesture

Coupling animation with gesture augments anthropomorphic agent believability. A gestural agent is different from an agent that is merely animated. Animation is movement; gestures are movements that contain information [19]. When gestures are effectively employed they are seldom noticed. They fade into the custom of human-human interaction but result in an overall naturalness of communication. “Gestures are so much a part of human communication that we seldom notice them” [19]. An anthropomorphic agent that exhibits gestures timed with speech greatly increases its believability. Actors that only talk are much less evocative than those that communicate with their bodies. Such acting will undoubtedly contribute to the illusion of life that may ultimately be used to persuade consumers, or worse, to dupe them.

### 3.1.3 Natural Language Communication

Natural language communication includes receiving and responding in natural language (either text or voice or perhaps a mix-and-match of the two<sup>6</sup>) and understanding natural language. Note that just because an agent receives and responds in natural language does not necessarily mean it understands it [29].

The illusion created by an anthropomorphic agent that uses natural language — this includes an agent with no physical embodiment that is anthropomorphic *because* it uses natural language — is a very strong one. Like gesture, natural language communication derives its power to convince from its roots in human-human communication. Agents without natural language understanding can still appear intelligent when they use natural language [29]. The illusion of intelligence can be so strong that in some circumstances people may be fooled into thinking a natural language agent is a real person [9].

When natural language communication is combined with gesture of various degrees, the illusion is even sturdier [5][33]. Such agents are convincing because of the apparent intelligence with which they synchronize gesture and speech. The goal of building one of these gesture-language agents [5] is to prove the opposite of a common warning about agent design: that building realistic agents unduly raises user expectations. The counter-argument is that the creation of a realistic gesture-language agent might prove to lower the difficulty in human-computer interaction, since humans are used to speaking and gesturing with humans, not cartoons [5]. If a competent conversational agent can be built, clearly it will have erected a fortified illusion of intelligence and life that could mislead consumers.

The natural language communication ability is arguably the strongest of all anthropomorphic qualities. Agents could be built to take full advantage of this capability to the point that humans fail to realize the agent is a limited computer program. Consumers may be lured into trusting advice from agents in their decisions for online transactions.

### 3.1.4 Emotion

An agent that can effectively convey appropriate emotional responses greatly augments the illusion of life that may mislead consumers. Joseph Bates writes, “The emotionless character is lifeless. It is a machine” [2]. Emotions are something we find at the heart of what it means to be human. Our emotions define us. They define our relation to others and our mental and physical well being. We think something so central to humanity could not be replicated in all its complexity inside a machine. However, it is not emotion that must be replicated, it is the *appearance* of emotion [11]. An agent that appears to respond emotionally will be treated as an emotional being, and thus interpreted by our minds as a life form. The illusion of life will be strengthened immensely by an agent that uses emotion with any kind of proficiency.

### 3.1.5 Personality

Though some research efforts (e.g., [23]) have tried to define personality rigorously, for our purposes personality can be treated as the summation of the characteristics above. Personality is conveyed in everything from a character’s movement [16] to its

choice of language [23] to its emotion [2]. Thus, personality is really a holistic package of consistent attributes all working together to contribute to an overall impression for the user.

We know when an agent has an attractive personality much like we know a character in a book or film has an attractive personality. And though our real world experiences are littered with interesting characters, our virtual worlds lack them [11]. The potential exists for people to become obsessed with an attractive virtual personality in much the same way that people become captivated by personalities in other media like television and film. A consumer who interacts regularly with an anthropomorphic agent might grow fond of that agent’s personality and be more likely to trust the agent, potentially with ill effects.

The strength of anthropomorphic agents lies in the illusion of life. The illusion of life is created when agents leverage those things with which we identify as “human.” We interact with such agents on a social level as though they are competent peers. Our placement of such agents in e-commerce will rely on the strength of these interactions for success. But there is potential for danger precisely because of the depth and intensity of human-agent interaction.

## 3.2 Human Tendencies in Human-Agent Interaction

Human tendencies exacerbate dangers of placing agents in e-commerce. Specifically, three attributes of human psychology make interactions with lifelike anthropomorphic agents potentially hazardous. As Jaron Lanier observes, “the only difference between an autonomous ‘agent’ program and a non-autonomous ‘editor/filter’ program is in the psychology of the human user. You change yourself in order to make the agent look smart” [20].

### 3.2.1 The Tendency to Personify

Personification takes place in many areas of life, not just with anthropomorphic agents. We name our cars, boats, and other vehicles. Anyone who has programmed for long hours on a computer has suspected the computer of sabotaging what should have been a perfectly working program. The pet rock phenomenon was an illustration of the human tendency to personify.

A lot of attention has been given to research showing that we treat computers and other media much like we treat other people, that is, socially [27]. Though in our minds we know that a computer is nothing more than a machine, we interact with it and regard it as though it were more. This evidence is even stronger if the machine is visibly anthropomorphic: “Give anything eyes and a mouth ... and personality responses follow” [27]. Brenda Laurel argues that our tendency to personify computers is because they *behave*; they have predispositions to behave in certain ways that we come to expect [22].

The result of our tendency to personify is a lower threshold for a suspension of disbelief. The agent designer need only approximate believable life, as humans are all too ready to invite a character into the realm of the living. Even crude stick figures are enough to elicit a social response from us [27]. If an agent effectively employs all the techniques in the previous section we are hopeless to resist the illusion — not that we *should* resist if the agent has our best interests at heart. But an agent that sells us goods or negotiates contracts might have an advantage over our

<sup>6</sup> For example, Extempo agents receive typed natural language but respond with text-to-speech synthesized voice. See [www.extempo.com](http://www.extempo.com).

human psyche by the nature of their design. The illusion could be exploited to our financial detriment.

### 3.2.2 The Tendency to Trust Specialists

Humans trust specialists far more than non-specialists when interacting in unfamiliar domains. This trust transfers to “specialist” computers and agents: “If media claim specialization, they’ll be better liked and appear more competent ... Ethical considerations aside, there is little doubt that claims to specialization work” [27]. Just as we trust advice from human specialists more than advice from well-meaning friends, we highly trust advice from media that claims specialty.

Agents acting as online sales personnel or domain-specific allies present the façade of specialization. We are more likely to believe a Nascar racing saleswoman agent [11] when purchasing a car than an unspecialized animal cartoon. Product endorsements by specialist anthropomorphic agents, as in the case of their human counterparts, achieve greater persuasiveness by exploiting the human tendency to trust.

### 3.2.3 The Tendency to Form Relationships

Relationships formed with anthropomorphic agents are a source of potential harm for humans. This harm may be financial or even emotional.

Humans are relational. They define themselves by their relationships with their family, friends, coworkers, and the divine. They are susceptible to forming relationships even with things that cannot form a relationship back, such as a favorite teddy bear [20].

Evidence of the enormous tendency for humans to form fervent relationships with anthropomorphic agents is not hard to find.<sup>7</sup> The public’s response to Bandai’s Tamagotchi virtual pet proved the depths of devotion between human and machine. The psychological trauma of the deaths of Tamagotchi pets on their owners has been reported [26], and vast virtual graveyards swell with flowing eulogies about the glorious life of these agents.<sup>8</sup> The success of PF Magic’s Petz collection is another testimony to the human tendency to attribute life to anthropomorphic agents and become attached to them, even form meaningful relationships with them. The formation of strong relationships between agents and humans has been raised as a point of ethical consideration, as we continue to “muddy the distinction between users being amused, or assisted ... and users creating an emotional attachment of some kind with the embodied image that the lifeless agent projects” [7].

The combination of anthropomorphic agent qualities and human tendencies result in possibly alarming e-commerce interactions.

## 4. LOOKING AHEAD TO AGENTS IN E-COMMERCE

Agent technology is improving rapidly and agents are on the brink of flooding the virtual workplace. At least four companies are now

<sup>7</sup>A somewhat humorous example of a human male attempting to form a sexual relationship with text-based “female” agent Julia can be found in [9].

<sup>8</sup> See, for example, [www.mirskyland.com/tamagot.htm](http://www.mirskyland.com/tamagot.htm), [www.d-3.com/deadpet/](http://www.d-3.com/deadpet/), [www.geocities.com/Tokyo/Flats/6337/](http://www.geocities.com/Tokyo/Flats/6337/)

commercially producing agents to populate corporate web sites and form virtual sales forces.<sup>9</sup> Soon we will visit an e-commerce site and find a familiar face waiting to greet us who remembers our name and with whom we have an ongoing relationship.

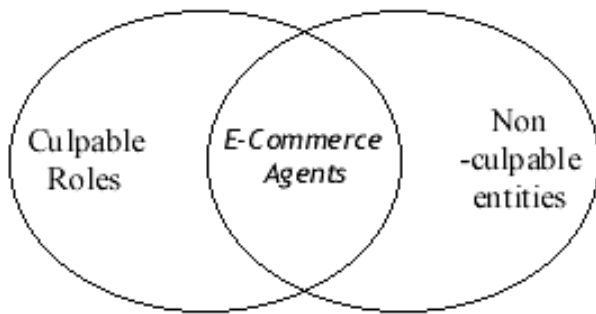
## 4.1 Move Over, Human! Agents in Human Roles

The rosy picture of an online agent-mediated future is excitedly circulating among agent designers and e-commerce entrepreneurs. But little thought has been given to the social and legal ramifications of such an arrangement. When anthropomorphic agents begin jobs as salespersons, service reps, guides, and hosts, they will not be mere user-interface widgets akin to buttons or toolbars. Buttons, toolbars, and other UI components are direct extensions of the user: they respond directly to a user’s actions and only for as long as the user acts upon them. Anthropomorphic agents, on the other hand, are not acted upon but are actors themselves. *When deployed, they will fill roles until now exclusively owned by human beings, from salespersons to dispensers of advice to sympathetic companions.* This is a revolution beyond computing. The social, economic, and legal spheres have never seen anything like it. Whether the effect on these domains is positive or negative will be jointly in the hands of the agent designers and policymakers.

The reason for concern is plain: Anthropomorphic agents will attempt to interact with people in a manner that emulates human-human interaction but will fail to achieve this in every case. For reasons discussed in the previous section, we know that this interaction is quite convincing, subconsciously if not consciously. What happens, however, when the agent cannot maintain the illusion on which the humanlike interaction depends? Who will be held culpable if the agent makes a mistake? What are some of the consequences of botched agent-human interactions? Will agents be employed to persuade people or defraud them? What rules will govern the acceptable and unacceptable in online agent behavior? Might users become addicted to relationships formed with online agents as some have with Tamagotchi virtual pets? Are there dangers if they do? These are only some of the many questions that arise out of the new way in which agents will fill human roles.

Anthropomorphic agents in human roles are filling what we might call *culpable roles*. These are roles that require an assumption of responsibility and often a duty to a client or to the public at large. If a human actor in a culpable role behaves negligently then that actor may be held liable. But e-commerce agents will soon be in these roles, and clearly *they* cannot be held liable. The result is a conundrum of accountability that will likely resolve with the agent designer and/or the person who configured the application held responsible. Users will probably be absolved of responsibility since they violated no duty during an interaction with an agent in which the agent made a vital mistake or compelled them into taking some action.

<sup>9</sup> See corporations Artificial Life ([www.artificial-life.com](http://www.artificial-life.com)), Extempo ([www.extempo.com](http://www.extempo.com)), Inago ([www.inago.com](http://www.inago.com)), and NetSage ([www.netsage.com](http://www.netsage.com)).



**Figure 1. The revolutionary merger of culpable roles with non-culpable entities at the point of anthropomorphic e-commerce agents**

It is not hard to imagine circumstances in which an online agent filling a culpable role brings liability to its designers or employers. For example, if Astrology Ally is an anthropomorphic agent designed to foster strong relationships with users by offering advice and remembering such things as a user's name, birthday, astrological information, likes and dislikes, and significant life details, then the user may become emotionally attached to the agent. A web site might charge users to receive Ally's investment advice: "the stars hold for you a great windfall in the near future." Users who followed such advice to their detriment — for example by investing large sums of money that was subsequently lost — might claim the agent was negligently designed. The agent's designers might be held liable for scripting it to give advice outside its plausible range of knowledge, especially if such advice benefited the agent's company in some way [3].

Another possibility is an agent that sells online goods. Perhaps such an agent is used to sell a high-technology gadget about which there is a great deal of intricate knowledge. If the user asks the agent detailed questions that the agent does not fully understand, the agent may invoke a non-comprehension (non-comp) procedure which involves an Eliza-like turnaround of the question or a change of subject [35], a common technique of non-player characters in MUDs. Such incomplete or inadequate — and potentially misleading — information for the user could result in a misinformed expenditure of money.

Skeptics might argue that humans should not be so easily fooled by an obviously limited computer character, but if the research is right, we know that humans are caught up in the illusion of life and respond socially as though the agent were competent and capable [27]. If users are left to fend for themselves in an online marketplace of misleading and defrauding agents, then e-commerce will suffer. We need to make consumers feel relaxed in the knowledge that the agents with which they interact are restricted in what they can say and do.

The law has always revolved around the assumption that people fill culpable roles; not just *any* people, but people that can be held responsible for their actions [6]. When anthropomorphic agents fill roles in online malls they will be filling culpable roles. Their designers must be prepared to assume responsibility for their agent employees, just as parents assume responsibility for their minor children.

## 4.2 Persuasion and E-Commerce Agent Ethics

Anthropomorphic e-commerce agents will be vehicles for persuasion because of their contact with humans and the convincing manner in which they operate. The illusion of life can be used by businesses to influence people's attitudes and behaviors. At this early stage it is impossible to envision all the persuasive capacities in which anthropomorphic agents will be employed. We must begin asking questions now as to the appropriate boundaries, if any, that should be placed on agent persuasion.

Boundaries will be crucial when agents are used in advertising. As Jaron Lanier writes, "If info-consumers see the world through agent's eyes, then advertising will transform into the art of controlling agents, through bribing, hacking, whatever" [20]. Advertisers will undoubtedly use anthropomorphic agents to persuade users to buy certain things or try new products since we know that anthropomorphic agents have a strong social and emotional effect on people. The extent to which we permit advertising anthropomorphic agents to persuade humans must be determined.

Concerns with advertising anthropomorphic agents are heightened when we remember that agents are often adaptive [15]. Adaptive agents can model a user's preferences and track a user's patterns. The agent then passes the user profile back to advertisers. While uses of agents like this may be legal, human-agent interaction in e-commerce will suffer if users have a fear of agents snooping on behalf of advertisers.

The new field of captology studies persuasive computing and is relevant to persuasive anthropomorphic agents in e-commerce. Captologists recognize the potential dangers with persuasive anthropomorphic agents. In fact, the field considers technologies that are "social actors" as one of its principal components [8]. Captologists consider social actors as vehicles for persuasion because of the way social actors directly inherit the intentions of their human creators. Researchers in this new field also rely on the finding that users respond socially to anthropomorphic agents, which heightens the potential for persuasion.

Captology contains a strong impetus to consider the ethics of persuasive computing technology [3][8]. Unethical persuasion in e-commerce becomes conning, and anthropomorphic agents are prime candidates for it.<sup>10</sup> Captologists have examined ethics in persuasive computing and recognize that in many cases of ethically questionable persuasive technology, "companies stand to gain profit or information, while individuals stand to lose money, privacy, or freedom" [8]. The charge for high-tech designers is to base their designs in defensible ethical standards. Captologists urge designers to avoid deception, respect privacy, and enhance personal freedom [8]. Future ethical standards that grow out of this field will help to inform anthropomorphic agent design ethics.

<sup>10</sup> We are all familiar with the stereotype of the sleazy used-car salesman masterfully pushing a lemon to unsuspecting buyers. Envisioning a sales agent of the same bent is not difficult.

## 5. REDUCING THE DANGER IN ANTHROPOMORPHIC AGENT DESIGN

We offer five straightforward anthropomorphic agent design suggestions for limiting threats to the electronic commerce consumer.

**Create transparent agents.** An agent should not appear to the user to be a black box from which apparently intelligent thoughts and behaviors come [13][14]. Jaron Lanier pointedly defines an agent as “a program that conceals a haphazard personality profile of a user from that user” [20] and contends that users concede their autonomy when interacting with an “opaque” agent. Instead agents should reveal their intentions and be required to explain their actions whenever asked [36]. Users should have some idea of what is going on inside the agent’s “mind” and what the agent knows about them. This will lessen the numbing effects of the illusion of life and empower users in their interactions with agents.

**Create humble agents.** A salesman who admits that he does not know the answer is admirable. Anthropomorphic sales agents should not follow the precedent set by Eliza [35], namely turning around an unanswerable question so as to maintain the illusion of life.<sup>11</sup> Rather, agents should inform users as to the boundaries of their abilities. Even further, agents should *encourage* users to explore the limits of their capabilities [25]. An agent’s honest disclosure by agents of its own limitations will equip and empower users.

**Avoid unnecessary realism.** We disagree with the view expressed in [5] that anthropomorphic agent design should strive for realism. There is no evidence that realism results in a better overall user-experience, and there is some evidence to the contrary [34]. Many agent designers advocate using non-realistic depictions so as to lower user-expectations [36] and in e-commerce this will be particularly salient. A good example of an effective yet simple level of anthropomorphism is Pattie Maes’ Maxims agent [24].

**Carefully consider agent-mediated persuasion.** Ever since the serpent persuaded Eve in the garden, persuaders have been ethically suspect [3]. While it will be tempting for designers to leverage the illusion of life to change user attitudes and behavior, such practices may result in user injury and designer liability. Intentions to persuade must be weighed carefully with a developed ethics of persuasive technology. Ethical systems such as that presented in [3] provide for helpful forethought.

**Facilitate user goals.** Users will enjoy interacting with agents if their goals are realized. Goal-facilitation is a directive of HCI and should not be forgotten just because the interface contains an autonomous character [36]. Users should still hold the definitive judgment on e-commerce anthropomorphic agents, and maintaining the illusion of life should be secondary to user safety.

## 6. CONCLUSION

The future of anthropomorphic agents in electronic commerce at present rests in the hands of pioneering anthropomorphic agent designers.

---

<sup>11</sup> This was not the goal in the creation of Eliza, but it may have set a precedent for future agent dialogue.

Following the five suggestions presented above is of limited effectiveness in the long run. No simple list could empower agent designers to guide their agent development efforts ethically and legally. Without a doubt there will be circumstances that challenge even these user-friendly maxims.

Agents will be commercially successful if they aid users in accomplishing their goals — not by sniffing information or manipulating emotional attachment. For example, the designer must ask *why* she is adding the ability to create a user profile. If she is building an agent to track users’ interests, would the possibilities for improved interaction and service be enough that the user himself would want that information collected? On one hand, it can make for a more personal online experience. On the other, it can allow advertisers to solicit intimately a user who has never heard of them before. Many issues will be double-edged like this; consequently, user welfare must remain the golden trophy.

What is needed is a framework for *thinking* about motivations in agent design. To put her best face forward, the agent designer must evaluate features in terms of user welfare. The user must be treated as an end and not a means to profit or information.

## 7. REFERENCES

- [1] Bates, Joseph. The Nature of Characters in Interactive Worlds and The Oz Project. In Loeffler, Carl Eugene (ed.) *Virtual Realities: Anthology of Industry and Culture*. Van Nostrand Reinhold, New York, 1993.
- [2] Bates, Joseph. The Role of Emotion in Believable Agents. *Communications of the ACM* 37, 7 (July 1994), 122-125.
- [3] Berdichevsky, Daniel and Neunschwander, Erik. Toward and Ethics of Persuasive Technology. *Communications of the ACM* 42, 5 (May 1999), 51-58.
- [4] Blumberg, Bruce. Old Tricks, New Dogs: Ethology and Interactive Creatures. MIT Media Lab Ph.D. dissertation, 1997.
- [5] Cassell, J. and Bickmore, T. and Billinghurst, M. and Campbell, L. and Chang, K. and Vilhjálmsón H. and Yan, H. Embodiment in Conversational Interfaces: Rea. *Proceedings of ACM SIGCHI '99*. Pittsburgh, PA, 1999, 520-527.
- [6] Duff, R. A. *Intention, Agency, and Criminal Liability*. Basil Blackwell, Cambridge, MA, 1990.
- [7] Elliott, Clark and Brzezinski, Jacek. Autonomous Agents as Synthetic Characters. *AI Magazine*, Summer 1998, 13-30.
- [8] Fogg, B. J. Persuasive Computers: Perspectives and Research Directions. *Proceedings of ACM SIGCHI '98*. ACM Press, Los Angeles, 1998, 225-232.
- [9] Foner, Leonard N. What’s an Agent, Anyway? *Agents Memo 93-01*. Available from the Agents Group, MIT Media Lab, Cambridge, MA, 1993. Available at <http://foner.www.media.mit.edu/people/foner/Reports/Julia/Agents--Julia.pdf>.
- [10] Gearino, G. D. Living Large in a Mall World. *News and Observer*. Raleigh, N.C., October 3, 1999, at D1.

- [11] Hayes-Roth, Barbara and Doyle, Patrick. Animate Characters. *Autonomous Agents and Multi-Agent Systems 1*, 1998, 195-230.
- [12] Hayes-Roth, Barbara and Sincoff, Erik and Brownston, Lee and Huard, Ruth and Lent, Brian. Directed Improvisation. *Knowledge Systems Laboratory Report No. KSL 94-61*. Stanford University, Stanford, CA, 1994. Available at [http://www-ksl.stanford.edu/KSL\\_Abstracts/KSL-94-61.html](http://www-ksl.stanford.edu/KSL_Abstracts/KSL-94-61.html).
- [13] Heckman, Carey and Roetter, Alexander. Designing Government Agents for Constitutional Compliance. *Proceedings of the Third International Conference on Autonomous Agents (Agents '99)*. ACM Press, Seattle, 1999.
- [14] Heckman, Carey and Wobbrock, Jacob O. Liability for Autonomous Agent Design. *Autonomous Agents and Multi-Agent Systems 2* (1), March 1999, 87-103. This article initially published as source [15] was extended to consider notions of transparent agents.
- [15] Heckman, Carey and Wobbrock, Jacob O. Liability for Autonomous Agent Design. *Proceedings of the Second International Conference on Autonomous Agents (Agents '98)*. ACM Press, Minneapolis, 1998, 392-399.
- [16] Jones, Chuck. *Chuck Amuck: The Life and Times of an Animated Cartoonist*. Avon Books, New York, 1990.
- [17] Kant, Immanuel. *Ethical Philosophy*. Hackett Publishing Company, Indianapolis, 1983. Originally published in 1785 entitled *Grounding of the Metaphysics of Morals*.
- [18] Krochmal, Mo. Online Shoppers Frustrated by Confusing Sites. *TechWeb News*, September 28, 1999.
- [19] Kurtenbach, Gordon and Hulteen, Eric A. Gestures in Human-Computer Communication. In B. Laurel (ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, Mass., 1990.
- [20] Lanier, Jaron. Agents of Alienation. *Interactions*. July 1995, 66-72.
- [21] Lasseter, John. Principles of Traditional Animation Applied to 3D Computer Animation. *Proceedings of ACM SIGGRAPH '87*. Anaheim, CA, July 1987, 35-44.
- [22] Laurel, Brenda. Interface Agents: Metaphors with Character. In B. Laurel (ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, Mass., 1990.
- [23] Loyall, Bryan A. and Bates, Joseph. Personality-Rich Believable Agents that Use Language. *Proceedings of the First International Conference on Autonomous Agents (Agents '97)*, ACM Press, Marina del Rey, CA, 1997, 106-113.
- [24] Maes, Pattie. Agents that Reduce Work and Information Overload. *Communications of the ACM 37*, 7 (July 1994), 31-40, 146.
- [25] Murray, Janet. Flat and Round Believable Agents. *Working Notes from the AAAI Spring Symposium Series on Believable Agents*, Stanford University, March 19-20, 1994, 65-68.
- [26] Neumann, Peter G. *comp.risks* vol. 19, issues 20, 36, 37. Illustrative Risks to the Public in the Use of Computer Systems and Related Technology. SRI International, Menlo Park, CA, January 1999.
- [27] Reeves, Byron and Nass, Clifford. *The Media Equation: How People Treat Computers, Televisions, and New Media Like Real People and Places*. Cambridge University Press, New York, 1996.
- [28] Rousseau, Daniel and Hayes-Roth, Barbara. Personality in Synthetic Agents. *Knowledge Systems Laboratory Report No. KSL 96-21*. Stanford University, Stanford, CA, 1996. Available at [http://www-ksl.stanford.edu/KSL\\_Abstracts/KSL-96-21.html](http://www-ksl.stanford.edu/KSL_Abstracts/KSL-96-21.html).
- [29] Schmandt, Chris. Illusion in the Interface. In B. Laurel (ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, Reading, Mass., 1990.
- [30] Sengers, Phoebe. Do the Thing Right: An Architecture for Action-Expression. *Proceedings of the Second International Conference on Autonomous Agents (Agents '98)*. ACM Press, Minneapolis, MN, 1998, 24-31.
- [31] Shneiderman, Ben and Maes, Pattie. Direct Manipulation vs. Interface Agents. *Interactions*. November/December 1997, 44-61.
- [32] Spool, Jared M. and Scanlon, Tara and Schroeder, Will and Snyder, Carolyn and DeAngelo, Terri. *Web Site Usability: A Designer's Guide*. Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1999.
- [33] Thórisson, Kristinn R. Real-Time Decision Making in Multimodal Face-to-Face Communication. *Proceedings of the Second International Conference on Autonomous Agents (Agents '98)*, ACM Press, Minneapolis, MN, 1998, 16-23.
- [34] Walker, Janet and Sproull, Lee and Subramani, R. Using a Human Face in an Interface. *Proceedings of ACM SIGCHI '94*. Boston, Mass, 1994, 85-91.
- [35] Weizenbaum, J. ELIZA: A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM 9*, 1 (January 1966), 36-45.
- [36] Wexelblat, Alan and Maes, Pattie. Issues for Software Agent UI. Unpublished manuscript by the *Software Agents Group, MIT Media Lab*, 1997. Available at <http://wex.www.media.mit.edu/people/wex/agent-ui-paper/agent-ui.htm>. Used with permission from the authors.