Takayama, L. (2011). Perspectives on agency: Interacting with and through personal robots.
In Zacarias, M. & Oliveira, J. V. (Eds.), Human-Computer Interaction: The Agency Perspective. Springer.

Perspectives on agency

Interacting with and through personal robots

Leila Takayama[1]

Abstract Personal robots present opportunities for understanding the ways that people perceive agency—both in-the-moment and reflectively. Autonomous and interactive personal robots allow us to explore how people come to perceive agency of non-human agents. Remote presence and tele-operation systems are expanding our understandings of how people interact through robots, incorporating these systems into their own sense of agency. As such, robotics can inform our understanding of both robotic agency and human agency.

Introduction

What does it mean for something to have agency? Some would say that even mundane objects have agency (Latour, 1992) while others would say that objects do not have agency. While it is possible to argue at length about the ontological status of an entity's agency, it is also possible to define agency as something that is perceived. Regardless of the absolute status of an entity's agency, it is our *perceptions* of agency that influence how we behave. This chapter focuses upon how we perceive agency and why those perceptions matter for our interactions with and through robots.

As a human-robot interaction researcher, it is very common for people I meet to spontaneously share with me their Roomba (robotic vacuum cleaner) stories. Some Roomba owners perceive their Roombas as agentic objects, whereas others perceive their Roombas as just plain machines. A Roomba owner reported feeling hesitant to return "Spot" for maintenance because the owner did not want to get a replacement Roomba back from the company; similarly, other Roomba owners exhibit protective behaviors, worrying about clearing away their household contents that "it could get choked on or stuck on," (Sung, Guo, Grinter, &

[1] Willow Garage, 68 Willow Road, Menlo Park, California, USA

Christensen, 2007). In contrast, other people do not assign names to their Roombas and do not form social relationships or orientations toward them; seven out of fourteen Roomba owners did not form social relationships with the Roomba in one study (Forlizzi & DiSalvo, 2006) and 572 out of 760 messages did not contain any descriptions of intimacy felt toward Roombas in another study (Sung, et al., 2007). In the case of a Roomba, it may reflectively seem like an agentic member of the family or it may seem like nothing more than an appliance. More complex vacuuming robots that map the room before vacuuming with pre-planned straight lines (Neato) may be perceived as more agentic than a Roomba, which randomly wanders until it hits obstacles; such behaviors may even make these robotic vacuum cleaners seem to be more intentional (Dennett, 1987).

There are a couple of concepts that are useful for untangling this seemingly contradictory set of responses—*in-the-moment* vs. *reflective* perspectives. An "in-the-moment" perspective refers to one's most immediate (sometimes visceral) sense in a situation, as when people say, "At the time, it seemed as if…" In contrast to "in-the-moment" perspectives, a "reflective" perspective refers to one's sense of a situation upon more distanced cogitation and consideration.

In-the-moment, a robotic vacuum cleaner may seem more like an agentic object as it moves around your house and your pets interact with it (as observed by Forlizzi, 2007), chasing it, riding on it, or barking at it. In the case of the iRobot Create, which can be teleoperated, the robot may seem more like a part of yourself in-the-moment when you drive it around your house—similar to demonstrations of self-extension demonstrated with Legos Mindstorm robots (Groom, Takayama, Ochi, & Nass, 2009) and rubber hands (Botvinick & Cohen, 1998). We reflectively believe that these objects are not actually parts of ourselves, but it can feel quite differently in the moment of interaction.

A critical dimension of agency is the perspective from which something seems to have agency. Neglecting to separate reflective perspectives from in-the-moment perspectives of agency is one of the major sources of confusion when people talk and write about anthropomorphism, ethopoeia (Nass, Steuer, Tauber, & Reeder, 1993), and computers as social actors (Nass & Moon, 2000; Reeves & Nass, 1996). People deny interacting with computational systems as if they were people and yet they respond to computers in many ways that are remarkably similar to how they respond to people (Reeves & Nass, 1996, pp. 26-27). This is not surprising, given that people often have little or no direct introspective access to the existence of the stimulus, to their responses, or the connection between the stimulus and their responses (Nisbett, 1977). One explanation for why people might respond to computers in ways they respond to people is that these people are responding mindlessly (Nass & Moon, 2000) rather than in a reflective manner. This issue becomes increasingly important when computational agents take on more embodied forms as in the case of many personal robots.

To grapple with this disconnect between what people consciously perceive and how they respond to stimuli that they may not consciously perceive, this chapter presents an analytical approach to understanding these phenomena. Those two perspectives are *in-the-moment* perceptions of agency and *reflective* perceptions of agency (Takayama, et al., 2009).

In the following sections we will present these different perspectives on agency as they apply to human-robot interaction. First, we will explicate in-the-moment vs. reflective perceptions of agency. Second, we will discuss robots as they are perceived in-the-moment as being agentic (e.g., autonomous robots)—interacting *with* robots. Third, we will discuss robots as they are perceived in-the-moment as being a part of one's sense of self (e.g., invisible-in-use robots)—interacting *through* robots. Each of these section will address the concepts that underlie those in-the-moment experiences and will ground those concepts in actual robots in use today. Finally, we will discuss what these perspectives on agency mean for the research and design of personal robots.

In-the-moment vs. Reflective Perspectives on Agency

In-the-moment perspectives consist of what people perceive and do when they are inside of a given situation, whereas reflective perspectives consist of what people perceive when they sit back to consider the situation from a more distanced perspective. See Figure 1 for a depictive description of these two different perspectives. The difference between these two situations is the location of the eye, which represents the perspective from which one looks upon the human-robot situation.

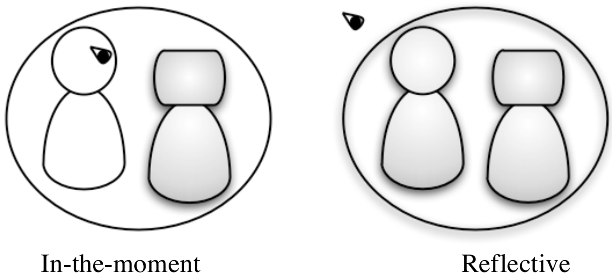

Figure 1. In-the-moment vs. Reflective perspectives on human-robot interaction.

Without making claims about what is objectively true or what is objectively real, it is still possible to distinguish between what is believed *reflectively* and what is perceived *in-the-moment*. In this sense, agency exists in the eye of the beholder. Although perceptions are often dismissed in favor of "objective" measures in many sciences, this is not reasonable when studying such agentic objects (Takayama, et al., 2009), where objective realities are contestable and subjective realities are the motivating bases for real, in-the-moment interaction with the agentic objects. Artificial intelligence researchers came up against this same problem of evaluating the success—sufficient intelligence—of artificial intelligence systems. As a result, the most famous measure of success became the Turing Test, which relied upon an observer's inability to distinguish a person from a computer in textual communication (Turing, 1950). This test leverages perceived similarity in favor of other seemingly objective measures of intelligence. Perceptions are not second best to objective measures; in fact, it is perceptions and subjective realities that people make judgments and act upon (as demonstrated by Kahneman & Tversky, 1974).

The other important distinction between the two perspectives is the cognitive processes that they engage. In-the-moment perceptions of agency are largely shaped by bottom-up perceptual processes, evoking very immediate responses. It is no surprise that the Wizard of Oz (Fleming, 1939) used a very large and animated image with a very loud, booming voice to project to his minions. These highly evocative cues of size, motion, and volume are useful for grabbing human attention; arguably, humans have evolved automatic responses to potential threats and opportunities in much less man-made situations than the ones we live in today (Nass & Gong, 2000). Just as Tinbergen's (1951) mother birds were easily cued to automatically roll anything with brown speckles into their nests, humans also have a propensity toward particular cues that seem to trigger automatic interpersonal social responses; this includes interactive computers that use language to communicatively trigger human-like responses from users (Nass, Steuer, & Tauber, 1994). In contrast, reflective perceptions are more often shaped by top-down processes because of the nature of reflective thought. The truth value of reflection perceptions is irrelevant to this situation. What matters is that the reflection perceptions are more consciously constructed.

Related to the distinction between in-the-moment and reflective perspectives are first vs. second-line reasoning, Heidegger's ready-at-hand vs. present-at-hand perspectives, and the Elaboration Likelihood Model (ELM). First vs. second-line reasoning about agency among humans, robots, and computers, which focused on how people predict subsequent behaviors (Levin & Saylor, 2009). Heidegger's notions of ready-at-hand vs. present-at-hand align with in-the-moment vs. reflective perspectives; he uses the example of the carpenter's hammer to demonstrate how a carpenter can use the hammer to pound nails (ready-at-hand) or the carpenter can sit back and reflect upon the feel of the hammer in his hand, the hammer's weight, etc. (present-at-hand) (Heidegger, 1992). The ELM is a model of communication that predicts different levels and types of message processing, depending upon whether the person engaged in focused, thoughtful processing (central route) or unfocused, mindless processing (peripheral route) (Petty, Cacioppo, & Schumann, 1983). All of these theories touch upon the notion that people perceive and think differently when they are more or less reflective about the situation.

These two perspectives on agency may be juxtaposed upon each other in a two-dimensional conceptual space depicted in Figure 2.

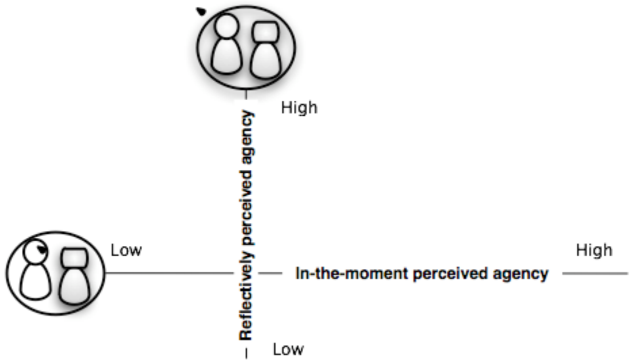

Figure 2. Two-dimensional space of in-the-moment vs. perceived agency.

The vertical dimension represents *reflective perception* of perceived agency. The horizontal dimension represents the *in-the-moment perception* of perceived agency. At the highest level of perceived agency, the entity seems to have its own needs, desires, and goals, e.g., a person. At a middle level of perceived agency, the entity does not seem to have agency of its own, but may seem to have potential affordances for the observer, e.g., a tool. At the lowest level of perceived agency, the entity not only has no agency, but is also incorporated into the observer's own agency and embodied *I can* (Leder, 1990), e.g., tool incorporated into one's embodied experience. The origin point of this space indicates the boundary between where the agentic object seems to be incorporated into part of oneself (X<0) and where agentic object seems to be separate entities (X>0). It also indicates where we reflectively believe that agentic objects are a part of oneself (Y<0) and where agentic objects are separate entities (Y>0). To ground this conceptual space presented in Figure 2, we present nine concrete examples, explaining each region of the space in Figures 3a-3c.

Each of the nine points within this two-dimensional space is an example of an entity that would fall in that particular combination of reflective and in-the-moment perceived agency. These are instances along a continuous dimensions of agency, nominally labeled as “low” to “high.”

Starting with the simplest cases (Figure 3a), I perceive other people such as yourself as highly agentic entities both in-the-moment and reflectively. I perceive inert robots as mediocre in agency both in-the-moment and reflectively. I perceive my foot as part of myself both in-the-moment and reflectively.

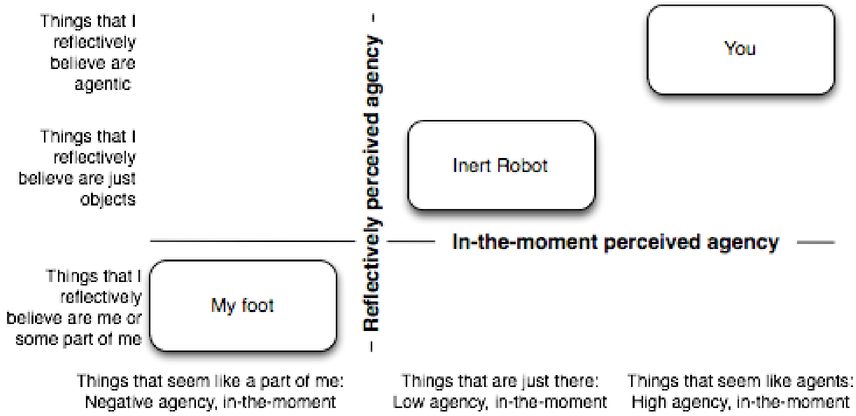

Figure 3a. Matched in-the-moment and reflective perceptions of agency (X=Y).

Slightly more complex cases include those points in the space where things are perceived as *more* agentic in-the-moment than they seem upon further reflection (Figure 3b). When my foot falls asleep it feels like an object hanging off of my leg (e.g., Leder, 1990, p. 85) but I reflectively believe it as part of me. When I see a surprising surveillance video feed of myself or hear a voice recording of myself (Holzman & Rousey, 1966), it feels like a different person in-the-moment though I can step back and perceive it as me, reflectively. When autonomous robots elicit social responses from people (Nass & Moon, 2000), they are being perceived in-the-moment as highly agentic entities even though people reflectively know they are merely computational machines.

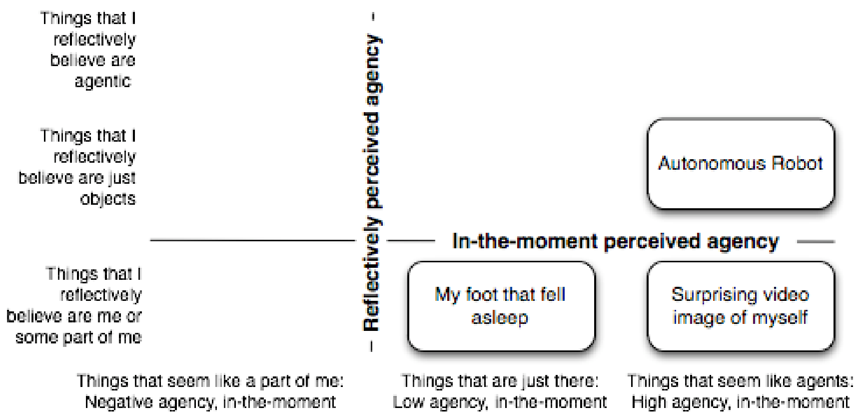

Figure 3b. In-the-moment perceptions of agency that are greater than reflective perceptions of agency (X>Y).

Another set of more complex cases include those points in the space where things are perceived as *less* agentic in-the-moment than they seem upon further reflection (Figure 3c). The service worker who discreetly sees to my every need may seem like a tool in-the-moment, but I know that this person is an autonomous individual, reflectively. The expert ballroom dancing partner engages in the joint activity (Clark, 1996) of dancing with me as if we were a single entity in-the-moment though we perceive each other as separate individuals, reflectively. The expert surgeon, who uses his teleoperated surgical device (e.g., the DaVinci (Guthart & Salisbury, 2000)) experiences the robot in-the-moment as part of himself, but believes reflectively that the robot is just another machine in the hospital.

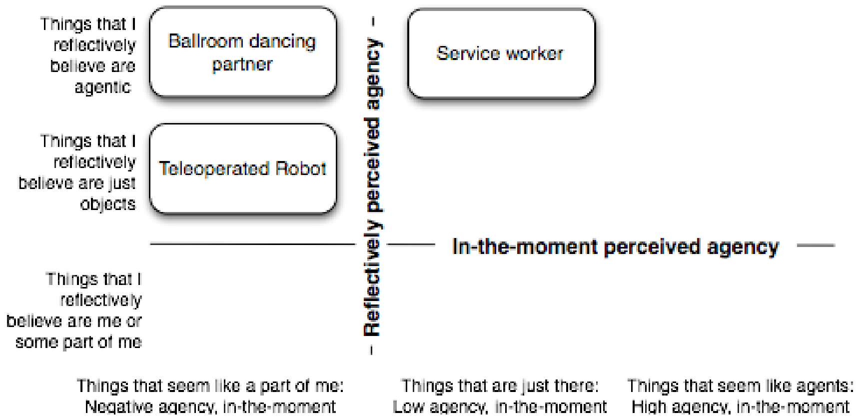

Figure 3c. In-the-moment perceptions of agency that are less than reflective perceptions of agency (X>Y).

It is precisely these in-the-moment perceptions of agency that are necessary for human interaction with agentic objects. As objects become increasingly endowed with computational "smarts" (Figure 3b), they become increasingly perceived in-the-moment as agents in their own right. Other fields such as teleoperation and ubiquitous computing leverage the computational power to make tools becoming part of the user's phenomenological self (Figure 3c), thereby becoming increasingly perceived in-the-moment as part of the user's own agency.

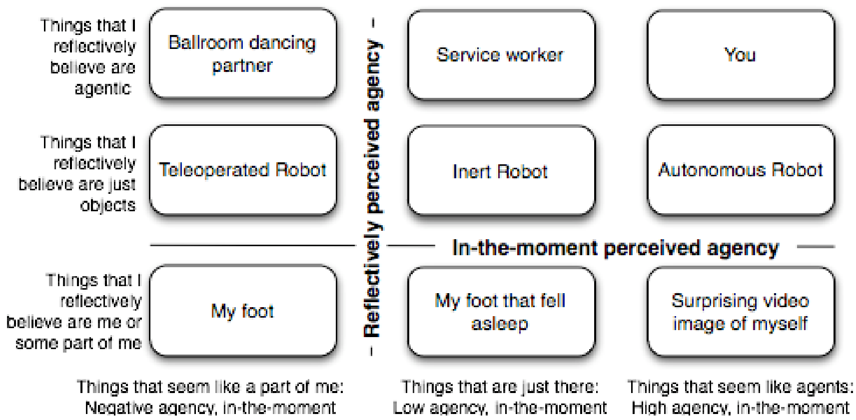

Figure 4. Levels of in-the-moment and reflective agency with example instances.

There are some notable hybrids of these two perspectives, including mindfulness (Varela, Thompson, & Rosch, 1991) and reflective design (Schon, 1983, 1992). The idea of these ways of being is to remain reflective while being in-the-moment and active. They recommend “a change in the nature of reflection from an abstract, disembodied activity to an embodied (mindful), open-ended reflection" (Varela, et al., 1991, p. 27). Schon would call this reflection-in-action as opposed to post-hoc reflection-on-action. While these are provocative and interesting ideas, they also rely upon the assumption that reflective and in-the-moment perspectives are most often completely separate in time.

Interacting with robots

Interacting *with* robots is one example of interacting with agentic objects, i.e., systems that seem to have their own agency. Examples of interacting with robots include playing rock-paper-scissors with Nico (Short, Hart, Vu, & Scassellati, 2010), chatting with Roboceptionist (Lee, Kiesler, & Forlizzi, 2010), or receiving a beer from PR2. (See Figure 5.)

Figure 5. Personal Robot (PR2) delivering a beer to Jon.

In this real life example, PR2 is behaving in a manner that is very similar to how a waiter might serve a drink to a customer. Despite the fact that Jon was one of the developers of the autonomous code that runs on the PR2 that enables it to deliver the drink, it can still feel like PR2 is a very agentic robot in-the-moment of interaction, receiving a bottle of beer from the robot. This is just one example of the many ways that personal robots are being designed to take on human-like roles, which imbue them with a sense of in-the-moment agency.

Agentic objects

Agentic objects (Takayama, 2009) are those entities that are perceived in-the-moment as if they were agentic despite the likely reflective perception that they are just objects—hardly agentic at all. As computational power seeps out of mainframes and personal computers into everyday objects, we are increasingly faced with agentic objects in our everyday lives. Agentic objects include complex artificial intelligence systems such as robots and "smart" spaces, but also simpler systems such as commonplace automatic doors, trashcans, sink faucets, and staplers that behave as though they have their own agency.

It may seem strange to attribute agency to non-human entities, but it is actually quite common. Agency is perceived in the relationships between entities (Knappett, 2002). Both infants and adults have been found "to treat novel self-moving objects as though they have both perception/attention, communicative abilities and goals if they either look like an agent (i.e., have a face) or behave in specific ways (e.g., are contingently interactive with other known agents)" (S. C. Johnson, 2003, p. 557). In a series of experiments, Johnson had a confederate interact with a non-human object—a beach ball size, fuzzy blob that could move, flash lights, and beep—and measure gaze following behaviors (i.e., seeing whether an infant looks at the "eyes" of the agent and looks in the direction of the object's "gaze" direction), which is an important step for engaging in joint attention with other agents. Adults denied that they believed the objects were "true agents," but they still responded to those objects in-the-moment as though they had agency, following the object's "gaze." Self-movement alone was not enough to elicit such responses from people though; this only happens when the object's behaviors are contingent with other people's behaviors.

As a commonplace example of an agentic object, we turn to automatic doors that, more often than not, "*discriminate* against very little and very old persons… and everyone with packages, which usually means… working or lower-middle-class employees" (Latour posing as J. Johnson, 1988). Such automata (i.e., self-moving machines) have been with us since at least the European Renaissance, if not earlier (e.g., the Temple at Delphi); these automata were primarily objects to be observed from afar, e.g., Vaucanson's defecating duck in 1738, though some were intended to interact with people, e.g., Wolgang von Kempelen's speaking machine in 1791, which met with limited success (Riskin, 2003). In this world of expanding sensor capabilities, growing computational power, and rapidly shrinking devices, we are encountering an increasing diversity of automata that are "smart" (as described by Gershenfeld (1999)).

How are we to make sense of these encounters with agentic objects? Furthermore, how are we to design for such encounters and their inevitable breakdowns? People seem to have an innate ability to respond to agentic objects in ways that are very similar to interacting with other humans (Nass & Moon, 2000; Reeves & Nass, 1996). This may be part of our evolutionary inheritance (Nass & Gong, 2000) or part of our use of pretense and imagination in learning and play (Harris, 2000; Walton, 1990). Many of us grew up with dolls, stuffed animals, or fictional characters in books and on television. Playing cops and robbers or holding pretend tea parties relies upon our ability to engage in joint pretenses with others (Harris, 2000). Now that objects are becoming increasingly agentic, they may be evoking new forms of joint pretense based upon existing forms of play and theater.

This goes beyond simply anthropomorphically interpreting the behavior of agentic objects (e.g., Bassili, 1976; Heider & Simmel, 1944) to the more important step of interacting with agentic objects. Automatic doors aim to open in much the same way doorman opened them, creating a similar perceptual experience for proximate pedestrians and creating a similar functional experience of revealing an entryway into a previously closed building (Ju & Takayama, 2008). When automatic doors open for people, those people accept or reject the invitation to enter the building. They do not merely talk about the door welcoming them into the building; they actually interact with it in that way—accepting and rejecting invitations to enter the doorway and feeling welcomed or offended, depending upon the behavior of the door.

Agentic objects are like Peirce's representamen in that their existence depends upon the perspective of interpretants. The representamen, "stands for something, just as an ambassador stands for his country, represents it in a foreign country; just as a deputy represents his electors in an assembly" (Deledalle, 2000, p. 39). Likewise, agentic objects inherit both the authority and the responsibility of other humans and human institutions. Automatic teller machines (ATMs) inherit the authority and responsibility of the banking institutions to which they belong. Automatic ticketing kiosks at airport check-ins inherit the authority and responsibility of the airlines that run them. Parking garage ticketing machines give tickets, receive tickets, and process payments with the authority and responsibility of the company that runs the garage.

Agentic objects are also like simulations. They are not perceived both in-the-moment and reflectively as the same type of agent. In the moment, they are virtual ("as if") agents. They are not representations of other beings; they are now generated to represent themselves as a simulacra. "Simulation is no longer that of a territory, a referential being, or a substance. It is the generation by models of a real without origin or reality: a hyperreal," (Baudrillard, 1994, p. 1). This definition of simulation also plays into the dimension of perceived agency in-the-moment. As Baudrillard writes, "To dissimulate is to pretend not to have what one has. To simulate is to feign to have what one doesn't have" (p. 3). To be perceived in-the-moment as less agentic than when perceived reflectively is to dissimulate. A modest service-worker is dissimulating. To be perceived, in-the-moment, as more agentic than when perceived reflectively is to simulate. An automatic door is simulating the desire of a doorman to welcome you into a building.

Because many robotic agents are designed for more complex tasks than simply opening a door, these agents require varying degrees of assistance from people. As Sheridan (1992) wrote in his book on telerobotics, automation, and supervisory control, there are many degrees of automation, ranging from the human doing everything to the computer deciding and acting autonomously, ignoring the human (p. 358); there are many levels in between such as the computer offering a set of action alternatives to the human, suggesting one action to the human, giving the human a chance to override the computer's selection, and merely informing the human of the computer's selection. In this book (Sheridan, 1992), Bill Verplank illustrated several ways that computers can share (supporting or extending human) and trade (backing up or replacing human) control over a task.

When people and robots need to communicate, e.g., when deciding what action to take next, it is important for robot user interfaces and human-robot interaction design to enable effective communication. The user interface issues for robots are very similar to the issues for computers in human-computer interaction, but now the physical embodiment of the robot (DiSalvo, Gemperle, Forlizzi, & Kiesler, 2002; Powers & Kiesler, 2006), its physical behaviors (Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005; Wang, Lignos, Vatsal, & Scassellati, 2006), and its other behaviors (e.g., speech (Fisher & Lohse, 2007)) also influence how people perceive and form mental models how to communicate with the system.

Robotic agency

Gold analyzed human-robot interaction as an information pipeline that includes the human, the robot, and the environment (Gold, 2009); this analysis identifies the human-to-robot communication as the bottleneck (i.e., most limited throughput of information). While progress is still being made in natural language processing, gesture understanding, learning by demonstration, etc., it is also possible for robot user interfaces to compensate for the low communication bandwidth between people and robots by leveraging existing human-computer interaction models (e.g, point-and-click UIs, DTMF phone-based UIs) (Goodfellow, et al., 2010).

Much of ongoing human-robot interaction research focuses upon creating the perception that robots are agentic objects to be interacted with. Looking at the space of in-the-moment vs. reflective agency (Figure 4), it is obvious that one likely way to seem more agentic in-the-moment (i.e., move right in the perceived agency space) is to behave in more humanlike ways, particularly as more independent people (e.g., less like the ballroom dancing partner of the service worker and more like an autonomous individual).

To this end, many robotics researchers are working toward developing and/or displaying human-like capabilities in robots. Ishiguro has built several androids that look remarkably similar to particular people, including himself—Geminoid (Sakamoto, Kanda, Ono, Ishiguro, & Hagita, 2007). Scassellati has been working toward creating developmental behaviors seen in infants, learning to engage in joint attention (Scassellati, 1999). Mutlu has developed and demonstrated how familiar interpersonal eye gaze behaviors can help people to engage and recall more information when ASIMO tells them a stories (Mutlu, Forlizzi, & Hodgins, 2006) and how robots like Robovie might influence participant roles in group conversations (Mutlu, Shiwa, Kanda, Ishiguro, & Hagita, 2009).

Even simpler robots are being designed to use humanlike behaviors. RoCo, the robotic desktop computer, engaged in humanlike gesturing behaviors (Breazeal, Wang, & Picard, 2007). Hoffman has demonstrated how robots such as AUR, the robotic desk lamp, can learn to anticipate actions from people and respond accordingly, making human-robot joint activity more fluid and efficient (Hoffman & Breazeal, 2010). Keepon, the snowman-shaped beat bot (Michalowski, Sabanovic, & Kozima, 2007), was built to exhibit eye contact and joint attention to learn about how children attribute mental states to robots (Kozima, Nakagawa, & Yano, 2004).

Social robots (Breazeal, Gray, Hoffman, & Berlin, 2004; Hegel, Muhl, Wrede, Hielscher-Fastabend, & Sagerer, 2009) are an active and relevant area of research and development. For example, Kidd's weight-loss coach robot, Autom, is designed to be an agent that holds you accountable for the calories that you consume and the exercise you do; in a controlled 6-week comparison of reporting calories and activities to Autom vs. a computer running the same software as Autom vs. a paper-based report used by the Boston Medical Center, the researchers found that people track their calories and exercise for twice as long as with the other methods (Kidd & Breazeal, 2008). In collaboration with the human-computer interaction communities that focus on virtual agents and avatars, along with the artists and designers in spaces such as animation and character design, there is still much to be learned about how to most effectively create physical agentic objects such as sociable robots.

While having and exhibiting humanlike behaviors moves robots into the territory of being agentic objects, there are other aspects of robots that also push them toward being perceived as being more agentic. Notice that when you move from the ballroom dancing partner over to the server and from the server over to being an individual person (Figure 4), there is a shift toward exerting one's own agency independent of the observer's wishes. When a person disagrees with you or otherwise acts against your wishes, you experience their exertion of agency more readily than when the person goes along with everything you say and do. Similarly, when an agent pushes against you, you perceive their agency more readily. As such, we have been conducting research in how robots could disagree with people if and when necessary (Takayama, et al., 2009). Similarly, when a robot cheats in a competitive game, people become more engaged in the game and make stronger claims about the robot's mental states than when the robot does not cheat (Short, et al., 2010). Even if these robots have the same amount of inherent abilities, it is their exertion of agency against your own wishes that makes them seem more agentic in-the-moment.

Just as robots can be designed to be perceived and interacted with as agentic objects, they are also being designed to be used more like the carpenter's trusty hammer or the blind man's cane—that is, they can become invisible-in-use.

Interacting through robots

Interacting *through* robots involves using robots to interact with people and the rest of the world, including robotic technologies to improve one's own sense of agency. Examples of interacting through robots include using a DaVinci teleoperation robot to perform surgery (Guthart & Salisbury, 2000; Intuitive), walking with support from the Berkeley Lower Extremity Exoskeleton (BLEEX) (Kazerooni, 2005; Kazerooni, Chu, & Steger, 2007), using an assistive robotic arm to reach and grab objects (Tsui, Yanco, Kontak, & Beliveau, 2008), using robots to find survivors in urban search and rescue missions (Murphy, 2004), or using remote presence systems to telecommute to work (Lee & Takayama, In Press). (See Figure 6.)

Figure 6. Remote presence system at work, attending a project meeting.

In this real life example, Dallas is piloting a mobile remote presence system to attend a project meeting that is taking place in the hallway. Because Dallas lives over 1500 miles away from the office, he telecommutes to work through this system every day and has been doing so for over a year now. In fact, many of Dallas' coworkers refer to this system as "Dallasbot" because it is the primary way that they interact with him. This mobile remote presence system has become an extension of Dallas in the local workplace.

Invisible-in-use

People incorporate robotic technologies into their own sense of agency, in-the-moment, although they reflectively believe that those technologies are separate from their bodies and senses of "self."

This idea is not new. This is the experience of the functional cyborg (Chorost, 2006), who incorporates tools into her own sense of capabilities (e.g., I can go 65 miles per hour), but does not actually believe that the tool is a part of her own body (e.g., I step out of my car, leaving it parked on the street). Similarly, Dant wrote about the first-person driver-car experience as an assemblage of person and object that change a person's orientation toward the world and the norms for daily social life (Dant, 2004). More specifically, ecological psychologist James Gibson explained how one's "field of safe travel" is like a tongue that protrudes in front of a driver-car, being shaped by obstacles, and potential hazards. He writes that "the car tends to become, like any properly used tool, simply a sort of physical extension of the driver's body" (Reed & Jones, 1982, p. 135). Leder called this an "incorporation" of a tool into one's body (1990) once it becomes so familiar that using it become a tacit experience rather than a focused, conscious one. Philosopher Michael Polanyi called this the "tacit dimension" of human experience; by knowing something tacitly (Polanyi, 1964), we come to perceive the world through it (e.g., a probe) rather than perceive the thing itself (e.g., feeling the probe in your hand). Both anthropologist Gregory Bateson and philosopher Maurice Merleau-Ponty point to the example of the blind man's cane (Bateson, 2000; Merleau-Ponty, 1962) to explicate their perspectives on this same idea that one's tools can become so familiar that they become incorporated into one's first-person perspective upon the rest of the world.

These notions of becoming so familiar with one's tools that they become invisible-in-use have also been discussed in the contexts of ubiquitous computing (Heer & Khooshabeh, 2004) and human-robot interaction (Takayama, 2010). In this context of understanding agency, it is best to think of making robots invisible-in-use as improving a person's own sense of agency.

Human agency

One of the most obvious ways to design for making robots invisible-in-use (i.e., move left in the perceived agency space) is by learning from other analogs in perceived agency space (Figure 4).

The simplest example my foot; when my foot falls asleep (i.e., restricted circulation makes my foot feel like a dead weight rather than like a part of my body), it is malfunctioning. Being a reliably functional system is a critical aspect of being invisible-in-use. If the carpenter's hammer splintered every once in a while, it would not be invisible-in-use.

Another example of moving left in the perceived agency space comes from the example of the surprising video image of myself. This is a breakdown in which it as not immediately obvious to me that the video image reflected me, not someone else. Providing perceptually immediate feedback (e.g., no perceptible time delay) between one's input and the robot's response to that input helps to move the robot toward the left side of the perceived agency space, being invisible-in-use. Teleoperation of robots has been focusing on this goal of creating a sense that the robot is invisible-in-use. One application domain of focus for this research community has been minimally invasive surgery, i.e., performing surgery through very small incisions, ideally smaller than the incisions that human surgeons have to create, e.g., the Black Falcon (Madhani, Niemeyer, & Salisbury, 1998). Intuitive Surgical's teleoperation robot, the DaVinci (Guthart & Salisbury, 2000) is an example of a robotic system that provides very responsive actions and immediate feedback to users (surgeons), thereby providing surgeons with the sense that they are just doing surgery, not that they are operating a large robot in order to perform surgery. For legal reasons and to get FDA approval for this system, it was critical to leave autonomous behaviors out of the DaVinci; the surgeons take responsibility for their actions so their actions must be mapped directly to the robot's performance.

Moving to other examples in the perceived agency space, we find that we can learn from examples like the ballroom dancing partner, whose behaviors are so predictable and in synchrony with my own that it feels like we share a sense of agency. If a human and robotic system can act with that degree of synchrony toward the person's goals, then the robot is more likely to become invisible-in-use. One example robotic system that seems to aim to do this is the Berkeley Lower Extremity Exoskeleton (BLEEX), which aids people in carrying heavier loads than their normal bodies would be able to handle (Kazerooni, 2005). They started out by measuring the human walking cycle (Zoss, Kazerooni, & Chu, 2006) so that BLEEX could sense the person's current walking behaviors in order to respond to and support the person. This robotic system amplifies the user's abilities rather than directly mirroring the user's behavior.

Of course, there are many degrees and types of autonomy that can be added to robotic systems, including extending, relieving, backing up, and replacing people (Sheridan, 1992). Mobile remote presence systems such as the Personal Roving Presence (ProP) (Paulos & Canny, 1998), PEBBLES for hospitalized children to attend school (Fels, Waalen, Zhai, & Weiss, 2001), HP's BiReality (Jouppi, Iyer, Thomas, & Slayden, 2004), and Texai (Willow Garage, 2010) present opportunities to understand how people can interact with other people through robots. These mobile remote presence systems allow for varying degrees of autonomous behaviors and "direct" teleoperated behaviors; for example, collision avoidance systems can help the pilot to avoid hitting walls and door, but that type of autonomous behavior also takes away the control from the pilot. It is possible for autonomous behaviors to become invisible-in-use as demonstrated by anti-lock braking systems and power steering.

We can also learn about how robots can become invisible-in-use through empirical studies about how tools become invisible-in-use. Common invisible-in-use tools that people mentioned in these studies included cars, cell phones, computers, pens, and contact lenses. Throughout these examples, there were some themes such as reliability, predictability, consistency, familiarity, and sense of control (Takayama, In Press). By drawing from these examples and from analogous lessons learned from other people and objects in the perceived agency space, we can inform the design of robots that become invisible-in-use, thereby improving a person's own sense of agency.

Agency in the context of personal robotics

Agency is a critical concept for research and the design of personal robotic systems, particularly because robotics brings more autonomy, sensors, and actuation to the mix of elements already present in the field of human-computer interaction. The distinctions between in-the-moment and reflective perspectives are also important to consider in the design of personal robots. It might be the case that you care more about influencing how people engage in interactions with the robot, in which case it makes more sense to focus on in-the-moment perceptions.

However, it might be that you care more about how people sit back, think, and talk reflectively with their friends about the robot, in which case it makes more sense to focus on influencing their reflective perceptions.

In thinking through the design of robotic systems, particularly for end-users, consider whose agency matters most: Are you designing to improve a user's sense of her own agency or are you designing an agentic object for her to interact with? Based on your answer to that question, you can find inspiration and guidance from other domains in which interactions encourage more or less agency in-the-moment.

Drawing a line between invisible-in-use and agentic robotic systems is a simplification of the problem. With any given system, it is easy to shift between different levels of perceived agency just as it is easy for the carpenter to shift from perceiving his hammer as being present-at-hand (e.g., feeling its weight, temperature, texture) to perceiving it as being ready-at-hand (e.g., just pounding nails). The purpose of presenting this space of perceived agency (Figure 4) is simply to draw out a dimensional space that can be used to glean inspiration for figuring out how to push in-the-moment perceptions of agency toward the right (i.e., seemingly more agentic) or toward the left (i.e., seemingly less agentic or even invisible-in-use), particularly in the case of robotics, but also in the broader cases of human-computer interaction and design. There are some examples that blur the difference between an agent and being invisible-in-use, including butlers (Sohn, Ballagas, & Takayama, 2009), who are supposedly so good at anticipating and responding to our needs that it is almost as if they are not there at all. As with any selection of a single variable or dimension, there are going to be limitations to the model, but there can still be useful insights and lessons to be learned from simplifying the design space.

With these concepts of in-the-moment vs. reflective perspectives and invisible-in-use vs. agentic objects, it is possible to make sense of what was once seemingly odd behaviors of Roomba owners. Roomba owners who set their robots to vacuum during the workday when no one is home, do not talk to their Roombas, and simply treat it like any other appliance in the house are most likely perceiving the Roombas to be invisible-in-use—at least, when they aren't cleaning the brushes or emptying it out. Roomba owners who follow their robots around as they vacuum and encourage their pets to play with the robot are most likely perceiving it in-the-moment as being an agentic object. Reflectively, the story can be quite different. Some Roomba owners clearly feel comfortable with naming their robots, writing about their emotional attachments to their robots, etc. Some do not. Of course, some Rooma owners will tell you that they know and believe that it is merely a machine and yet they will coax and cheer on the Roomba when it seems to be struggling. Other Roomba owners will tell stories about how belligerently "Rosie" had behaved during the wee hours of the morning, banging on everyone's bedroom doors and yet show no emotional attachment or interest in the robot when you visit. These people are not weird, hypocritical, or in denial; they are simply exhibiting different behaviors at different times because they are engaging in different orientations toward the Roombas in-the-moment vs. reflectively.

With a bit of perspective on agency and how people can reasonably have different beliefs about agency at different points in time, we can shed some light on our understanding of how people interact with and think about personal robots and we can begin to appreciate just how much perceived agency matters.

References

Bassili, J. N. (1976). Temporal and spatial contingencies in the perception of social events. *Journal of Personality and Social Psychology, 33*, 680-685.

Bateson, G. (2000). *Steps to an Ecology of Mind*. Chicago: University of Chicago Press.

Baudrillard, J. (1994). *Simulacra and simulation* (S. F. Glaser, Trans.). Ann Arbor: The University of Michigan Press.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eye sees. *Nature, 391*, 756.

Breazeal, C., Gray, J., Hoffman, G., & Berlin, M. (2004). *Social robots: Beyond tools to partners*. Paper presented at the International Symposium in Robot and Human Interactive Communication (RO-MAN), Japan.

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). *Effects of nonverbal communication on efficiency and robustness in human-robot teamwork*. Paper presented at the International Conference on Intelligent Robots and Systems (IROS), Edmonton, Canada.

Breazeal, C., Wang, A., & Picard, R. (2007). *Experiments with a robotic computer: Body, affect and cognition interactions*. Paper presented at the International Conference on Human Robot Interaction (HRI), Washington DC, USA.

Chorost, M. (2006). *Rebuilt: My journey back to the hearing world*. Boston, MA: Mariner Books.

Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.

Dant, T. (2004). The driver-car. *Theory, Culture, and Society, 21*(4/5), 61-79.

Deledalle, G. (2000). *Charles S. Peirce's philosophy of signs: Essays in the comparative semiotics*. Bloomington, Indiana: Indiana University Press.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

DiSalvo, C., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). *All robots are not created equal: the design and perception of humanoid robot heads*. Paper presented at the Designing Interactive Systems (DIS), London, England.

Fels, D. I., Waalen, J. K., Zhai, S., & Weiss, P. L. (2001). *Telepresence under exceptional circumstances: Enriching the connection to school for sick children*. Paper presented at the Interaction, Japan.

Fisher, K., & Lohse, M. (2007). *Shaping naive users' models of robots' situation awareness*. Paper presented at the International Conference in Robot and Human Interactive Communication (RO-MAN), Jeju Island, Korea.

Fleming, V. (Writer). (1939). The Wizard of Oz. USA: MGM.

Forlizzi, J. (2007). *How robotic products become social products: An ethnographic study of cleaning in the home*. Paper presented at the International Conference on Human-Robot Interaction (HRI), Washington DC, USA.

Forlizzi, J., & DiSalvo, C. (2006). *Service robots in a domestic environment: A study of the Roomba vacuum in the home*. Paper presented at the International Conference on Human Robot Interaction (HRI), Salt Lake City, USA.

Gershenfeld, N. (1999). *When things start to think*. New York: Henry Holt and Company.

Gold, K. (2009). *An information pipeline model of human-robot interaction*. Paper presented at the International Conference on Human Robot Interaction (HRI), San Diego, USA.

Goodfellow, I. J., Koenig, N., Muja, M., Pantofaru, C., Sorokin, A., & Takayama, L. (2010). *Help me help you: Interfaces for personal robots*. Paper presented at the International Conference on Human Robot Interaction (HRI), Osaka, Japan.

Groom, V., Takayama, L., Ochi, P., & Nass, C. (2009). *I am my robot: The impact of robot-building and robot form on operators*. Paper presented at the International Conference on Human Robot Interaction (HRI), San Diego, USA.

Guthart, G. S., & Salisbury, J. K. J. (2000). *The Intuitive telesurgery system: Overview and application*. Paper presented at the International Conference on Robotics and Automation (ICRA), San Francisco, USA.

Harris, P. L. (2000). *The work of the imagination*. Oxford: Blackwell Publishers Ltd.

Heer, J., & Khooshabeh, P. (2004). *Seeing the invisible*. Paper presented at the International Conference in Advanced Visual Interfaces (AVI), Gallipoli, Italy.

Hegel, F., Muhl, C., Wrede, B., Hielscher-Fastabend, M., & Sagerer, G. (2009). *Understanding social robots*. Paper presented at the International Conferences on Advances in Computer-Human Interactions (ACHI), Cancun, Mexico.

Heidegger, M. (1992). *Basic Writings: From Being and Time (1927) to the Task of Thinking (1964)*. Harper Collins Publishers, New York, NY.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*(2), 243-259.

Hoffman, G., & Breazeal, C. (2010). Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots, 28*(4), 403-423.

Holzman, P. S., & Rousey, C. (1966). The voice as a percept. *Journal of Personality and Social Psychology, 4*(1), 79-86.

Intuitive. DaVinci Surgical System, 2010, from http://www.intuitivesurgical.com/products/davinci_surgicalsystem/

Johnson, J. (1988). Mixing humans and nonhumans together: The sociology of a door-closer. *Social problems, 35*(3), 298-310.

Johnson, S. C. (2003). Detecting agents. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 549-559.

Jouppi, N. P., Iyer, S., Thomas, S., & Slayden, A. (2004). *BiReality: Mutually-immersive telepresence*. Paper presented at the International Conference on Multimedia, New York, NY.

Ju, W., & Takayama, L. (2008). *Approachability: How people interpret automatic door movement as gesture*. Stanford University.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Kazerooni, H. (2005). Exoskeletons for human power augmentation. In B. Siciliano & O. Khatib (Eds.), *Robotics Handbook* (pp. 3459-3464): Springer-Verlag.

Kazerooni, H., Chu, A., & Steger, R. (2007). That which does not stabilize, will only make us stronger. *International Journal of Robotics Research, 26*(1), 75-89.

Kidd, C. D., & Breazeal, C. (2008). *Robots at home: Understanding long-term human-robot interaction*. Paper presented at the International Conference on Intelligent Robots and Systems (IROS), Nice, France.

Knappett, C. (2002). Photographs, skeuomorphs and marionettes: Some thoughts on mind, agency and object. *Journal of Material Culture, 7*(1), 79-177.

Kozima, H., Nakagawa, C., & Yano, H. (2004). Can a robot empathize with people? *Artificial Life and Robots, 8*(1), 83-88.

Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In W. Bijker & J. Law (Eds.), *Shaping Technology / Building Society: Studies in Sociotechnical Change* (pp. 225-258). Cambridge: MIT Press.

Leder, D. (1990). *The absent body*. Chicago: University of Chicago Press.

Lee, M. K., Kiesler, S., & Forlizzi, J. (2010). *Receptionist or information kiosk: How do people talk with a robot?* Paper presented at the ACM Conference on Computer Supported Collaborative Work (CSCW), Savannah, USA.

Lee, M. K. & Takayama, L. (In Press). *"Now I have a body": Uses and social norms of mobile remote present in the workplace*. Proceedings at the ACM Conference on Human Factors in Computing Systems (CHI), Vancouver, CA.

Levin, D. T., & Saylor, M. M. (2009). *Distinguishing defaults and second-line conceptualization in reasoning about humans, robots, and computers*. Paper presented at the International Conference on Human Robot Interaction (HRI), San Diego, USA.

Madhani, A. J., Niemeyer, G., & Salisbury, J. K. J. (1998). *The Black Falcon: A teleoperated surgical instrument for minimally invasive surgery*. Paper presented at the Intelligent Robotic Systems.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (C. Smith, Trans.). New Jersey: Routledge & Kegan Paul.

Michalowski, M. P., Sabanovic, S., & Kozima, H. (2007). *A dancing robot for rhythmic social interaction*. Paper presented at the International Conference on Human Robot Interaction (HRI), Washington DC, USA.

Murphy, R. R. (2004). Rescue robotics for homeland security. *Communications of the ACM, 47,* 66-68.

Mutlu, B., Forlizzi, J., & Hodgins, J. (2006). *A storytelling robot: Modeling and evaluations of human-like gaze behavior*. Paper presented at the International Conference on Humanoid Robots, Genova, Italy.

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). *Footing in human-robot conversations: How robots might shape participant roles using gaze cues*. Paper presented at the International Conference on Human Robot Interaction (HRI), San Diego, USA.

Nass, C., & Gong, L. (2000). Social Aspects of Speech Interfaces From an Evolutionary Perspective: Experimental Research and Design Implications. *Communications of the ACM*, *43*(9), 36-43.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103.

Nass, C., Steuer, J., & Tauber, E. R. (1994). *Computers as social actors*. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI), Boston, MA.

Nass, C., Steuer, J., Tauber, E. R., & Reeder, H. (1993). *Anthropomorphism, agency, & ethopoeia: Computers as social actors*. Paper presented at the Interact, Amsterdam, The Netherlands.

Neato. Neato XV-11 Vacuum Cleaning Robot, 2010, from http://www.neatorobotics.com/

Nisbett, R. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Paulos, E., & Canny, J. (1998). *PRoP: Personal roving presence*. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI), Los Angeles, USA.

Petty, R., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *The Journal of Consumer Research, 10*(2), 135-146.

Polanyi, M. (1964). *The Tacit Dimension*. Garden City, NY: Doubleday & Co.

Powers, A., & Kiesler, S. (2006). *The advisor robot: Tracing people's mental model from a robot's physical attributes*. Paper presented at the International Conference on Human Robot Interaction (HRI), Salt Lake City, Utah.

Reed, E., & Jones, R. (Eds.). (1982). *Reasons for realism: Selected essays of James J. Gibson*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

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/CSLI.

Riskin, J. (2003). The defecating duck, or, the ambiguous origins of artificial life. *Critical Inquiry, 29*(4), 599-633.

Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H., & Hagita, N. (2007). *Android as a telecommunication medium with a human-like presence*. Paper presented at the International Conference on Human Robot Interaction (HRI), Washington DC, USA.

Scassellati, B. (1999). Imitation and the mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot *Computation for metaphors, analogy, and agents* (pp. 176-195): Springer-Verlag.

Schon, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Schon, D. (1992). Designing as reflective conversation with the materials of a design situation. *Research in Engineering Design, 3*(3), 131-148.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. Cambridge, MA: MIT Press.

Short, E., Hart, J. W., Vu, M., & Scassellati, B. (2010). *No fair!! : An interaction with a cheating robot*. Paper presented at the International Conference on Human Robot Interaction (HRI), Osaka, Japan.

Sohn, T., Ballagas, R., & Takayama, L. (2009). *At your service: Using butlers as a model to overcome the mobile attention deficit*. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI), Boston, USA.

Sung, J.-Y., Guo, L., Grinter, R. E., & Christensen, H. I. (2007). *'My Roomba is Rambo': Intimate home appliances*. Paper presented at the International Conference on Ubiquitous Computing (Ubicomp), Innsbruck, Austria.

Takayama, L. (2009). *Making sense of agentic objects and teleoperation: In-the-moment vs. reflective perspectives*. Paper presented at the International Conference on Human Robot Interaction (HRI), San Diego, USA.

Takayama, L. (2010). *On making robots invisible-in-use*. Paper presented at the AISB Symposium on New Frontiers in Human Robot Interaction, Leicester, UK.

Takayama, L. (In Press). Toward making robots invisible-in-use: An exploration into invisible-in-use tools and agents. In K. Dautenhahn & J. Saunders (Eds.), *New Frontiers in Human-Robot Interaction*: John Benjamins.

Takayama, L., Groom, V., & Nass, C. (2009). *I'm sorry, Dave: I'm afraid I won't do that: Social aspects of human-agent conflict*. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI), Boston, USA.

Tinbergen, N. (1951). *The study of instinct*. London: Oxford University Press.

Tsui, K., Yanco, H., Kontak, D., & Beliveau, L. (2008). *Development and evaluation of a flexible interface for a wheelchair mounted robotic arm*. Paper presented at the International Conference on Human Robot Interaction (HRI), Amsterdam, The Netherlands.

Turing, A. (1950). Computing machinery and intelligence. *Mind, 59*, 433-460.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and the human experience*. Cambridge, Massachusetts: MIT Press.

Walton, K. L. (1990). *Mimesis as Make-believe: On the Foundations of the Representational Arts*. Cambridge, Mass: Harvard University Press.

Wang, E., Lignos, C., Vatsal, A., & Scassellati, B. (2006). *Effects of head movement on perceptions of humanoid robot behavior*. Paper presented at the International Conference on Human Robot Interaction (HRI), Salt Lake City, Utah.

Willow Garage. (2010). Texai Alpha Prototype, from http://www.willowgarage.com/pages/robots/texai/overview

Zoss, A., Kazerooni, H., & Chu, A. (2006). On the biomechanical design of the Berkeley Lower Extremity Exoskeleton (BLEEX). *Transactions on Mechatronics, 11*(2), 128-138.