How Humans Respond to Robots:

Building Public Policy through Good Design

Heather Knight, Robotics Institute, Carnegie Mellon University | July 2014

Introduction

Historically, robotics in industry meant automation, a field that asks how machines perform more effectively than humans. These days, new innovation highlights a very different design space: what people and robots can do better together. Instead of idolizing machines or disparaging their shortcomings, these human-machine partnerships acknowledge and build upon human capability. From autonomous cars reducing traffic accidents, to grandparents visiting their grandchildren by means of telepresence robots, these technologies will soon be part of our everyday lives and environments. What they have in common is the intent to support or empower the human partners with robotic capability and ultimately complement human objectives.

Human cultural response to robots has policy implications. Policy affects what we will and will not let robots do. It affects where we insist on human primacy and what sort of decisions we will delegate to machines. One current example of this is the ongoing campaign by Human Rights Watch for an international treaty to ban military robots with autonomous lethal firing power—to ensure that a human being remain "in the loop" in any lethal decision. No such robots currently exist, nor does any military have plans to deploy them, nor is it clear when robotic performance is inferior, or how it is different from human performance in lethal force situations. Yet the cultural aversion to robots with the power to pull the trigger on their own is such that the campaign has gained significant traction.

Cultural questions will become key on a domestic, civilian level too. Will people be comfortable getting in an airplane with no pilot, even if domestic passenger drones have a much better safety record than human piloted-commercial aviation? Will a patient be disconcerted or pleasantly surprised by a medical device that makes small talk, terrified or reassured by one that makes highly accurate incisions? Sociability, cultural background and technological stereotypes influence the answers to these questions.

My background is social robotics, a field that designs robots with behavior systems inspired by how humans communicate with each other. Social roboticists might incorporate posture, timing of motion, prosody of speech, or reaction to people and environments into a robot's behavioral repertoire to help communicate the robot's state or intentions. The benefit of such systems is that they enable bystanders and interaction partners to understand and interact with robots without prior training. This opens up new applications for embodied machines in our everyday lives—for example, guiding us to the right product at Home Depot.

Heather Knight *is a PhD candidate at Carnegie Mellon and founder of Marilyn Monrobot, which produces robot comedy performances and an annual Robot Film Festival. Her current research involves human-robot interaction, non-verbal machine communications and non-anthropomorphic social robots.*

My purpose in this paper is not to provide detailed policy recommendations but to describe a series of important choices we face in designing robots that people will actually want to use and engage with. Design considerations today can foreshadow policy choices in the future. Much of the current research into human-robotic teams seeks to explore plausible practical applications given improved technological knowhow and better social understandings. For now, these are pre-policy technical design challenges for collaborative robots that will, or could, have public policy implications down the road. But handling them well at the design phase may reduce policy pressures over time.

From driverless cars to semi-autonomous medical devices to things we have not even imagined yet, good decisions guiding the development of human-robotic partnerships can help avoid unnecessary policy friction over promising new technologies and help maximize human benefit. In this paper, I provide an overview of some of these pre-policy design considerations that, to the extent that we can think about smart social design now, may help us navigate public policy considerations in the future.

"Will people be comfortable getting on an airplane with no pilot?"

Human Cultural Response to Robots

If you are reading this paper, you are probably highly accustomed to being human. It might feel like nothing special. But after 12 years in robotics, with researchers celebrating when we manage to get robots to enact the simplest of humanlike behaviors, it has become clear to me how complex human actions are, and how impressive human capabilities are, from our eyesight to our emotive communication. Unlike robots, people are uniquely talented at adapting to novel or dynamic situations, such as acknowledging a new person entering the room while maintaining conversation with someone else. We can identify importance in a complex scene in contexts that machines find difficult, like seeing a path in a forest. And we can easily parse human or social significance, noticing that someone is smiling but is clearly blocking your entry, for example, or knowing without asking that a store is closed. We are also creative and sometimes do unpredictable things.

By contrast, robots perform best in highly constrained tasks—for example, looking for possible matches to the address you are typing into your navigation system within a few miles of your GPS coordinates. Their ability to search large amounts of data within those constraints, their design potential for unique sensing or physical capabilities–like taking a photograph or lifting a heavy object--and their ability to loop us into remote information and communications, are all examples of things we could not do without

help. Thus, machines enable people, but people also guide and provide the motivation for machines. Partnering the capabilities of people with those of machines enables innovation, improved application performance and exploration beyond what either partner could do individually.

To successfully complete such behavior systems, the field of social robotics adapts methodology from psychology and, in my recent work, entertainment. Human coworkers are not just useful colleagues; collaboration requires rapport and, ideally, pleasure in each other's company. Similarly, while machines with social capabilities may provide better efficiency and utility, charismatic machines could go beyond that, creating common value and enjoyment. My hope is that adapting techniques from acting training and collaborating with performers can provide additional methods to bootstrap this process. Thus, among the various case studies of collaborative robots detailed below, I will include insights from creating a robot comedian.

"...as social creatures, it is often our default behavior to anthropomorphize moving robots."

Robots do not require eyes, arms or legs for us to treat them like social agents. It turns out that we rapidly assess machine capabilities and personas instinctively, perhaps because machines have physical embodiments and frequently readable objectives. Sociability is our natural interface, to each other and to living creatures in general. As part of that innate behavior, we quickly seek to identify objects from agents. In fact, as social creatures, it is often our default behavior to anthropomorphize moving robots.

Animation is filled with anthropomorphic and non-anthropomorphic characters, from the Pixar lamp to the magic carpet in Aladdin. Neuroscientists have discovered that one key to our attribution of agency is goal-directed motion.[1] Heider and Simmel tested this theory with animations of simple shapes, and subjects easily attributed character-hood and thought to moving triangles.[2]

To help understand what distinguishes object behavior from agent behavior, imagine a falling leaf that weaves back and forth in the air following the laws of physics. Although it is in motion, that motion is not voluntary so we call the leaf an object. If a butterfly appears in the scene, however, and the leaf suddenly moves in close proximity to the butterfly, maintaining that proximity even as the butterfly continues to move, we would immediately say the leaf had "seen" the butterfly, and that the leaf was "following" it. In fact, neuroscientists have found that not attributing intentionality to similar examples of goal-directed behavior can be an indication of a social disorder.[3] Agency attribution is part of being human.

One implication of ascribing agency to machines is that we can bond with them regardless of the machine's innate experience, as the following examples demonstrate. In 2008, Cory Kidd completed a study with a robot intended to aid in fitness and weight loss goals, by providing a social presence with which study participants tracked their routines.[4] The robot made eye-contact (its only moving parts), vocalized its greetings and instructions, and had a touch-screen interface for data entry. Seeking to keep in good standing, it might have tried to re-engage participants by telling them how nice it was to see them again if they had not visited the robot in a few days. Its programming included a dynamic internal variable rating its relationship with its human partner.

When Kidd ran a study comparing how well participants tracked their habits, he compared three groups: those using pen and paper, touchscreen alone, or touchscreen with robots. While all participants in the first group (pen and paper) gave up before the six weeks were over, and only a few in the second (touch screen only) chose to extend the experiment to eight weeks when offered (though they had all completed the experiment), almost all those in the last group (robot with touch screen) completed the experiment and chose to extend the extra two weeks. In fact, with the exception of one participant who never turned his robot on, most in the third group named their robots and all used social descriptives like "he" or "she" during their interviews. One participant even avoided returning the study conductor's calls at the end of the study because she did not want to return her robot. With a degree of playfulness, they had treated these robots as characters and perhaps even bonded with them. The robots were certainly more successful at engaging them into completing their food and fitness journals than non-social technologies.

> *"With a degree of playfulness, they had treated these robots as characters and perhaps even bonded with them."*

Sharing traumatic experiences may also encourage bonding, as we have seen in soldiers that work with bomb-disposal robots.[5] In the field, these robots work with their human partners, putting themselves in harm's way to keep their partners from being in danger. After working together for an extended period, the soldier might feel that the robot has saved his life again and again. This is not just theoretical. It turns out that iRobot, the manufacturers of the Packbot bomb-disposal robots, have actually received boxes of shrapnel consisting of the robots' remains after an explosion with a note saying, "Can you fix it?" Upon offering to send a new robot to the unit, the soldiers say, "No, we want that one." That specific robot was the one they had shared experiences with, bonded with, and the one they did not want to "die."

Of course, people do not always bond with machines. A bad social design can be difficult to interpret, or off-putting instead of engaging. One handy rubric referenced by robot designers for the latter is the Uncanny Valley.[6] The concept is that making machines more humanlike is good up to a point, after which they become discomforting (creepy), until you achieve human likeness, which is the best design of all. The theoretical graph of the Uncanny Valley includes two lines, one curve for agents that are immobile (for example, a photograph of a dead person would be in the valley), and another curve with higher peaks and valleys for those that are moving (for example, a zombie is the moving version of that).
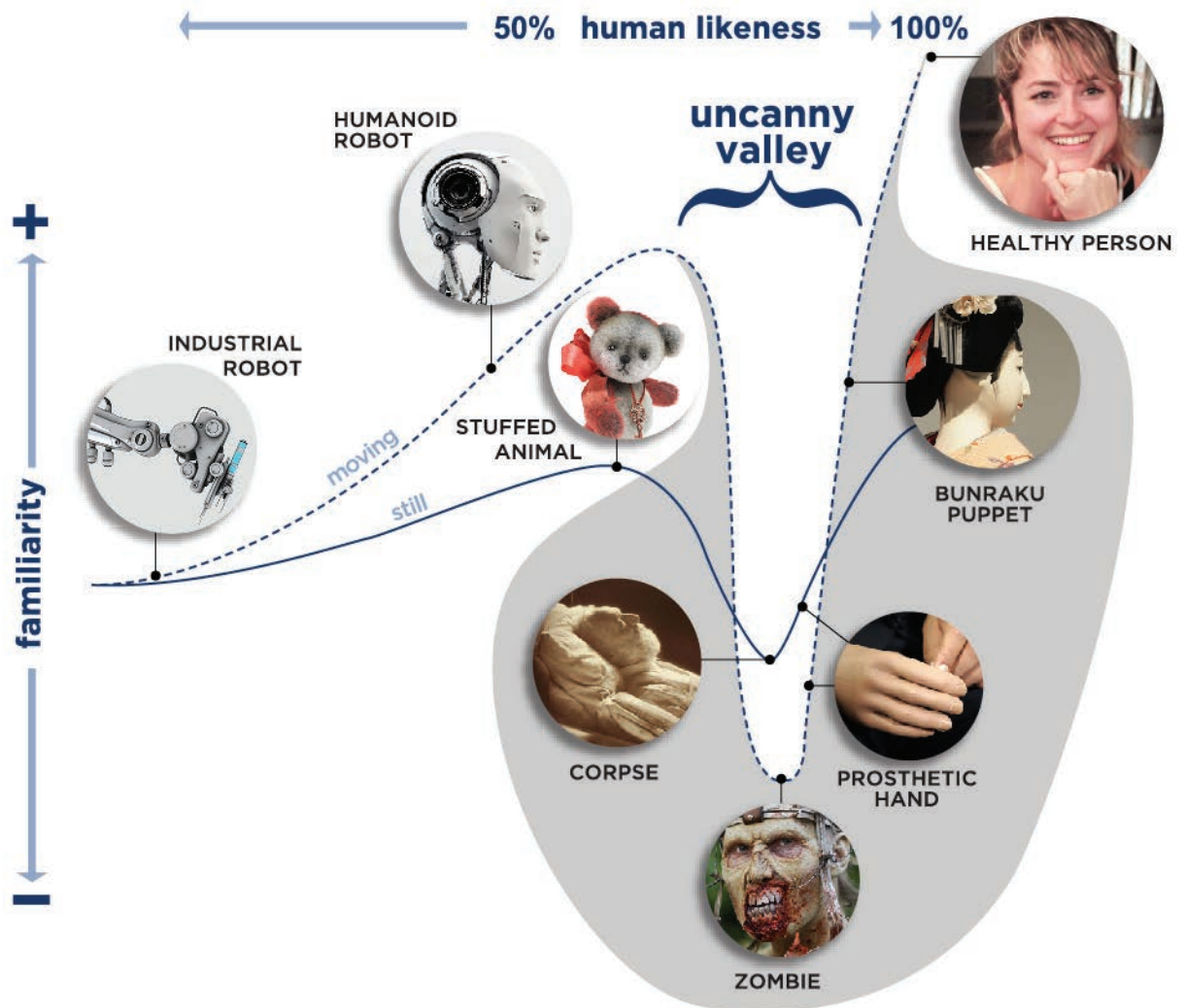
In my interpretation, part of the discomfort in people's response to robots with very humanlike designs is that their behaviors are not yet fully humanlike, and we are extremely familiar with what humanlike behavior should look like. Thus, the more humanlike a robot is, the higher a bar its behaviors must meet before we find its actions appropriate. The robotic toy Pleo makes use of this idea. It is supposed to be a baby dinosaur, an animal with which we are conveniently unfamiliar. This is a clever idea, because unlike possible robotic pets as dogs or cats, we have nothing to compare it against in evaluating its behaviors. In many cases, it can be similarly convenient to have more cartoonized or even non-anthropomorphic designs. There is no need for all robots to look, even somewhat, like people.

Cultural Variations in Response to Robots

"Big Brother no longer appears to be quite such a problem for many people."

Our expectations of robots and our response to their designs varies internationally; the Uncanny Valley curve has a different arc depending where you are. Certainly, our storytelling diverges greatly. In Japan, robots are cute and cuddly. People are apt to think about robotic pets. In the United States, by contrast, robots are scary. We tend to think of them as threatening.

Cultural response matters to people's willingness to adopt robotic systems. This is particularly important in areas of service, particularly caregiving services of one sort or another, where human comfort is both the goal and to some degree necessary for human cooperation in achieving that goal.

News reports about robot technologies in the United States frequently reference doomsday scenarios reminiscent of Terminator or RoboCop, even when the innovation is innocuous. My PhD advisor, Reid Simmons, jokes that roboticists should address such human fears by asking ourselves, "How prominent does the big red button need to be on the robots we sell?" Although there are also notable examples to the contrary (Wall-E, Johnny-5, C3PO), it is true that Hollywood likes to dramatize scary robots, at least some of the time (Skynet, HAL, Daleks, Cylons).

One explanation for the differing cultural response could be religious in origin. The roots of the Western Terminator complex may actually come from the predominance of monotheistic faiths in the West. If it is God's role to create humans, humans that create manlike machines could be seen as usurping the role of God, an act presumed to evince bad consequences. Regardless of current religious practice, such storytelling can permeate cultural expectations. We see this construct in Mary Shelley's story of Frankenstein, first published in 1818. A fictional scientist, Dr. Frankenstein, sews corpses

together then brings his super-creature to life in a lightning storm. Upon seeing it animated, he is horrified at the result, and in its abandonment by its creator, the creature turns to ill behavior. This sense of inevitability is cultural, not logical.

In Japan, by way of contrast, the early religious history is based on Shintoism. In Shinto animism, objects, animals, and people all share common "spirits," which naturally want to be in harmony.[7] Thus, there is no hierarchy of the species, and left to chance, the expectation is that the outcome of new technologies will complement human society. In the ever-popular Japanese cartoon series Astroboy, we find a very similar formation story to Frankenstein, but the story's cultural environment breeds an opposite conclusion. Astroboy is a robot created by a fictional Ministry of Science to replace the director's deceased son. Initially rejected by that parent figure, he joins a circus where he is rediscovered years later and becomes a superhero, saving society from human flaws.

"Social robots tap into our human need to connect with each other."

As one journalist writes, "Given that Japanese culture predisposes its members to look at robots as helpmates and equals imbued with something akin to the Western conception of a soul, while Americans view robots as dangerous and willful constructs who will eventually bring about the death of their makers, it should hardly surprise us that one nation favors their use in war while the other imagines them as benevolent companions suitable for assisting a rapidly aging and increasingly dependent population."[8] Our cultural underpinnings influence the media representations of robotics, and may influence the applications we target for development, but it does not mean there is any inevitability for robots to be good, bad or otherwise. Moreover, new storytelling can influence our cultural mores, one of the reasons why entertainment is important. Culture is always in flux.

Robot Design and Policymaking in the Information Age

Human cultural response to technology in general has been entering a rapid shift. Several decades ago, no one could have foreseen the comfort many of today's teenagers have with so much of their personal data appearing online. Big Brother no longer appears to be quite such a problem for many people. Cellphone providers might share our location data with other companies, the government might be able to read our emails, but in a world where social presence insists on constant Twitter updates and Instagram photos, people are constantly encouraged to be sharing this data with the world anyway.

The sexual revolution of the 1960s and 1970s did not last forever, and the pendulum may also swing back on data privacy, but reasoned aspects of the cultural shift in relation to technology will remain.

Social robots tap into our human need to connect with each other, and there are thus unique ethical and cultural considerations that arise in introducing such robots into society. The flipside of considering human bonding with machines is that robotic designers, and ultimately policymakers, may need to protect users from opportunities for social robots to replace or supplement healthy human contact or, more darkly, retard normal development. Think about a more extreme version of how video games are occasionally used by vulnerable populations (for example, the socially isolated or the depressed) as an escape that may or may not keep them from reengaging with other humans. Robot intelligence and simulated social behaviors are simplistic compared to the human equivalent. One cannot supplant the other, and protections should be in place to avoid asocial overreliance.

On the other hand, it is also possible to seek out ways of using these technologies to encourage human connection. As some autism researchers are investigating, social robots might help socially impaired people relate to others,[9] practicing empathetic behaviors as a stepping stone to normal human contact. In all, the ramifications of a robot's social role and capabilities should be taken into account when designing manufacturing guidelines and regulating consumer technologies.

In addition to encouraging positive applications for the technology and protecting the user, there may also ultimately arise the question of whether we should regulate the treatment of machines. This may seem like a ridiculous proposition today, but the more we regard a robot as a social presence, the more we seem to extend our concepts of right and wrong to our comportment toward them. In one study, experimenters had subjects run through a variety of team-building exercises together with six robotic dinosaurs, then handed them a hammer and asked the participants to destroy the robots. All of the participants refused.[10] In fact, the only way the study conductors could get the participants to damage any of the robots was to threaten to destroy all of the robots, unless the group hammered apart at least one. As Carnegie Mellon ethicist John Hooker once told our Robots Ethics class, while in theory there is not a moral negative to hurting a robot, if we regard that robot as a social entity, causing it damage reflects poorly on us. This is not dissimilar from discouraging young children from hurting ants, as we do not want such play behaviors to develop into biting other children at school.

Another effect of regarding robots (and other machines) as agents is that we react socially to their behaviors and requests. Some robots should be intentionally machine-like. In the 1960's, legend has it that the first automobile seat-belt detectors debuted with sound clips that could tell passengers to fasten their seat belts. Initially, owners celebrated the ability of their cars to talk, inviting the neighbors to see it and saying they were living in the future. After the novelty wore off, however, the idea of their car's giving them orders became not only irritating, but socially affronting. The personified voice of the car was trying to tell them what to do. Our reactions to social agents differ from our reactions to objects, and eventually car manufacturers sought to reduce the affront by changing the notification to a beeping sound instead.

Different human contexts can benefit from robots with more machine-like or more human-like characteristics. In an experiment where robots were theoretically to perform a variety of assistive tasks, participants were asked to select their preference for robot-like, mixed human-robot or human-like face.[11] In the context of personal grooming, such as bathing, most participants strongly preferred a system that acted as equipment only. Perhaps users were uncomfortable with camaraderie because of the personal nature of the task. On the other hand, when selecting a face to help the subject with an important informational task, such as where to invest the subject's money, the robotic face was chosen least, i.e., users preferred presence of humanlike characteristics. Younger participants selected the mixed human-robot face, and older participants generally selected the most humanlike face, perhaps because having human qualities made the robots seem more trustworthy. Understanding these kinds of interface expectations will influence the welcome and performance of any robot operating in a human context.

A Framework for Human-Robot Partnerships

To motivate design and policy considerations, I divide human-robot partnerships into three categories, each with short-term industrial or consumer applications: Telepresence Robots, Collaborative Robots, and Autonomous Vehicles. In the first category, humans give higher level commands to remote systems, such as remote-piloting in a search and rescue scenario. In the second, we introduce robots into shared human environments, such as our hospitals, workplaces, theme-parks or homes (for example, a hospital delivery robot that assists the nurses in bringing fresh linens and transporting clinical samples). In the third, people travel within machines with the ability to provide higher-level commands like destination or share control—for example, landing a plane after a flight takes place on auto-pilot.

For all of these categories, effective robots must have intuitive interfaces for human participation, clear communication of state, the ability to sense and interpret the behaviors of their human partners, and, of course, be capable of upholding their role in the shared task. We mine the examples within each category for design and policy considerations. The hope is to enable designs in which the split of human and machine capabilities enables users and has positive human impact, resulting in behaviors, performance and societal impact that go beyond what either partner could do alone.

Telepresence

Telepresence offers the ability for people to have sensing capabilities and physical presence in an environment where it is difficult, dangerous, or inconvenient for a person to travel. When I was working at NASA's Jet Propulsion Laboratory, the folks there liked to think of their spacecrafts and rovers as "extensions of the human senses." Other telepresence applications and developments include a news agency's getting a better view of a political event, a businessperson's skipping a long flight by attending a conference via robot, a scientist's gathering data, or a search-and-rescue team's leading survivors to safety.

Because of the distance, all of these tasks would require remote control interfaces and models for shared autonomy. The decision of how much computation and control should be delegated to the machine (shared autonomy) can exist on a sliding scale across different robots. Demanding higher local autonomy by the robot would be important when it needs rapid response and control, say, to maintain altitude, has unique knowledge of its local environment so it can, for example, orient to a sound, or is out of communications range. Assigning greater levels of control to the person (or team of people) can be a good choice when the task requires human expertise, the robot is in view or there are liability considerations. Even the Mars rovers have safeguards preventing them from completing their commands in the presence of an unexpected obstacle or danger--like, for example, the edge of a cliff--given that the operators would be too far away to communicate rapidly with the robots.

The autonomy breakdown is similar to the distinction in human bodies between reflexes and cognitive decisions. When you touch a hot burner, your spinal cord sends a rapid response to pull your hand back without consulting the brain. You don't think about it or make a decision to withdraw your hand. It just happens—a biological adaptation to safeguard the organism. Similarly, in situations where control loops would not occur fast enough,

communication goes down, or where we would want the pilot to oversee other robots, we can use shared autonomy to make use of the operator's knowledge of saliency and objective without weighing him or her down with the minutiae of reliable machine tasks, such as executing a known flight path.

The challenge of designing an interface for such a system is enabling a balance between human and machine capabilities that effectively accomplishes the task without overloading the operator or underachieving the objective. One way to enable this is to adapt the ratio of humans to robots. Ideally, human talents can complement and improve machine performance, but this only works if they don't have to micromanage the robots. To provide an example of an application with high robot autonomy, imagine a geologist that wants to use half a dozen flying robots to remap a fault line after earthquakes to better predict future damage. Using a team of robots, this scientist can safely explore a large swath of terrain, while the robots benefit from the scientist's knowledge of salience: what are the areas to map, say, or the particular land features posing danger of future collapse?

"Rather than missing an important meeting, a sick employee might log into the office robot and drive it to the conference room."

The human-to-robot ratio question has been an issue in military deployment of flying robots, often called drones. The operator's cognitive load will influence task performance and fatigue.

Designs that encourage simple control interfaces and appropriate autonomy breakdowns can help manage the pilot's cognitive load, enabling safety and better decision-making. Task complexity, required accuracy, and operator health might all be factors in regulating what the maximum assignments for each pilot should be. (There are also various social and policy issues in selecting autonomy designs this space, such as the ordinance in Colorado seeking to make shooting drones out the sky legal[12].)

In search-and-rescue tasks, a much lower level of robot autonomy might be desirable because of the potential complexity and danger. Operational design should reflect one overriding consideration: human safety. In current deployments, operators are often on site, using the robots to obtain aerial views that humans cannot easily duplicate. That means they could face danger themselves, so in certain environments, research teams have found success with three-person teams.[13] While the main pilot watches the point of view (video feed) of the robot and controls its motion, a second person watches the robot in the sky and tells the pilot about upcoming obstacles or goals that are out of robot view. A third acts as safety officer, making sure the first two do not become so engrossed in flying and watching the robot that they step into dangerous environments themselves. Standardizing safety regulations for such tasks, as such systems become more widely used, could save lives.

Another application area for telepresence robots is remote presence, in which a single individual can attend a remote event via robot. For example, rather than missing an important meeting, a sick employee might log into the office robot and drive it to the conference room. Once there, it might be more practical (and natural) for the robot to localize and orient toward whoever is currently speaking using its microphone array, extending the idea of local autonomy. Such behaviors might automatically keep the other speakers in view for the remote user, while displaying to the attendees his or her attention. While the remote user talks, the robot might automatically make "eye contact" with the other attendees by means of its ability to rotate.

Our social response to machines can be an asset to their impact and usability. In the above example, the physical presence of the robot and such natural motions could subconsciously impact the other attendees, resulting in more effective and impactful communication of the remote user's ideas. Furthermore, when the meeting is over, the remote user can chat with people as they exit the room or go with them as they visit the coffee machine down the hall. Such a machine could help display personality and presence in a more natural way than mere calling in or video-conferencing. Establishing trust, discussing ideas and getting along with each other is an essential part of productivity, and telepresence interfaces will benefit from incorporating that knowledge.

One policy concern of having such systems in the workplace is privacy, either of intellectual property discussed during the meeting or of the images captured of colleagues. Perhaps video data is deleted once transmitted and use of the robot is predicated on a binding agreement not to capture the video data remotely. In order for a telepresence system to maintain its social function of empowering the sick employee, the company and colleagues must know they are safe from misuse of the office robot's data, and firm policies and protections should be in place. Think of wiretapping laws.

The workplace is not the only potential user of remote presence systems. Researchers at Georgia Tech have begun to investigate the potential for use of telepresence robots by the elderly.[14] Whether by physical handicap or loss of perception, such as sight, losing one's driver's license can cause a traumatic loss of independence as one ages. One idea to counteract those feelings of isolation might be to enable an older person to attend events or visit his or her family by means of telepresence robot. During an exploratory study, researchers learned that while the elderly they surveyed had a very negative reaction to the idea that their children might visit them by robotic means, rather than coming in person, the idea of having a robot in their *children's* home that *they* could log into at will was very appealing.

The difference, of course, is that the former could increase their social isolation, while the latter example increases the user's sense of personal freedom. Some study participants expressed the desire to simply "take a walk outside," or attend an outdoor concert "in person." The trick here is to use such technologies in ways that protect the populations whom the technology is meant to support, or encourage the consumer use cases likely to have positive social impact.

Everyday Robots

Let us define everyday robots as those that work directly with people and share a common environment. From Rosie the Robot to remote-controlled automata at Walt Disney World, these systems can capture our imagination. They may also aid our elderly, assist workers, or provide a liaison between two people – like, for example, a robotic teddy bear designed to help bridge the gap between nurse and child in an intimidating hospital environment.

"In the right context, a robot acknowledging its mistakes with self-deprecation could make people like it better."

The physical and behavioral designs of such robots are usually specialized to a particular domain (robots do best in prescribed task scenarios), but they may be able to flexibly operate in that space with human assistance; for example, the Baxter robot is an industrial robot that can easily learn new tasks and uses social cues. Because of their close contact with people, the effectiveness of everyday robots will require analysis of social context, perception of human interaction partners, and generation of socially appropriate actions.

From robotic toys to robots on stage, there is a large market potential for robots in entertainment applications. Robots can be expensive; thus a high ratio of people to robots for these applications is likely. Entertainment can also provide a cultural benefit in helping shape what we imagine to be possible. The Disney theme parks have long incorporated mechanized character motions into park rides and attractions. More recently, they have begun to include remote-controlled or partially autonomous robots such as a robotic dinosaur, a friendly trash can, and a holographic turtle.

I may be biased, but there is also research value to placing robots in entertainment contexts. Think of a theater audience that changes from one evening to the next. With the right privacy protections for the data collected about the audience, the stage can provide a constrained environment for a robot to explore small variations iteratively from one performance to the next, using the audience members as data points for machine learning.[15] As I began to explore with my robot comedian, it may also be easier for robots to

interpret audience versus individual-person social behaviors because of the aggregate statistics (average movement, volume of laughter, and the like) and known social conventions (applause at the end of a scene).[16]

Even beyond machine learning algorithms, what we see in audience responses to robot entertainers may provide insight to other collaborative robots. In the right context, a robot acknowledging its mistakes with self-deprecation could make people like it better. In my experience with robot comedy, people love hearing a robot share a machine perspective: talking about its perception systems, the limitations of its processor speed, and battery life, and overheating motors bring a sense of reality to an interaction. The common value created may or may not invest the human partners in the robot, but it can definitely equip interaction partners with a better understanding of actual robot capabilities, limitations and current state—particularly if the information is delivered in a charismatic manner.

Another benefit of bringing performance methodology and collaborations into robotics research is what we can cross-apply from acting training. Not all robots talk. But even without verbalization, people will predict a robot's intention by its actions. Acting training helps connect agent motivation to physicality and gesture. Thus, adapting methodology across these domains could be particularly beneficial to the development of non-verbal robot communications, and also for creating consistent, sympathetic or customized personas for particular robot applications.

One of the benefits of robotic performance is that it creates greater understanding of human-robotic comfort levels—understanding that can then be applied elsewhere. Robots operating in shared spaces with humans benefit from learning our social patterns and conventions. A robot delivering samples in a hospital will not only need to navigate the hallways, they will need to know how to weave through people in a socially appropriate way.

In a study comparing hospital delivery robots operating on a surgery ward versus a maternity ward, the social context of its deployment changed the way people evaluated the robot's job performance.[17] While the surgeons and other workers in the former became frustrated with the robots getting even slightly in the way in the higher stress environment, the same robots in the maternity ward were rated to be very effective and likeable. Though its functionings were precisely the same, the same behavior system resulted in very different human response. This highlights the necessity of customizing behavior and function to social and cultural context.

Autonomous Vehicles

One might say that "autonomous" vehicles are already in operation at most airports. Pilots take care of takeoff and landing, but on the less challenging portion of the trip, for which people have a harder time maintaining alertness, an autopilot is often in charge unless something unusual happens. Japanese subway systems also run on a standard timetable with sensing for humans standing in the door. In both systems, there is the option for humans to take over, reset or override machine decisions, so ultimately these are shared autonomy systems with a sliding scale of human and machine decision-making.

I have heard people say that one of the reasons that flying robots have become so popular is that they have so little to collide with in the sky. In other words, machine perception is far from perfect, but if a system stays far away from the ground, there are few obstacles and it can rely minimally on human backup. But what happens in the case of autonomous cars navigating obstacle-rich environments surrounded by cars full of people? Such environments highlight the importance of design considerations that enable—and regulatory policies that require—such vehicles to learn, follow and communicate the rules of the road in a socially appropriate and effective fashion.

Autonomous cars have made rapid inroads over the past few years. Their immediate benefits would include safety and convenience for the human passenger; imagine not having to worry about finding a parking space while running an errand, because the car can park itself. Their habitual use would affect infrastructural changes, as parking lots could be further away from an event and traffic rules might be more universally followed. In some ways, semi-autonomous systems provide a clear shortcut for policymakers for liability considerations. By keeping a human in the loop, the fault in the case of bad decision-making leading to an accident becomes easier to assign.

What will become increasingly tricky, however, is the idea of changing the distribution of decision making, such that the vehicle is not just in charge of working mechanics (extracting energy from its fuel and transferring the steering wheel motion to wheel angle), but also for driving (deciding when to accelerate, or who goes first at an intersection). We already have cars with anti-lock brakes, cruise control, and distance-sensing features. The next generation of intelligent automobiles will shift the ratio of shared autonomy from human-centric to robot-centric. Passengers or a human conductor will provide higher-level directives (choosing the destination, requesting a

different route, or asking the car to stop abruptly when someone notices a restaurant they might like to try).

Vehicle technology should be designed to empower people. It might be tempting to look at the rising statistics of accidents due to texting while driving and ban humans from driving altogether, but partnering with technology might provide a better solution. If the human driver is distracted, a robotic system might be able smooth trajectories and maintain safety, much like a surgery robot can remove tremors of the human hand. The car could use pattern recognition techniques or even a built in breathalyzer to detect inebriation with a high probability and make sure the passengers inside make it home safely. Even without that technology, they might disable car function until they are in a better state to drive.

Additional safety considerations include the accuracy and failure modes for vehicle perception systems. They must meet high standards to make sure automated systems are actually supplying an overall benefit. Companies manufacturing the vehicles should be regulated as with any consumer technology, but customers may control local variables. If the passengers are in a rush, will they be able to turn up the car's aggression, creeping out into the intersection, even though that Honda Civic probably arrived a second or so before us?

With a user tweaking local variables or reprogramming certain driving strategies, who the technology's creator really is may become a bit of a moving target. Accidents will happen sometimes, and although cars cannot be sued in court, a manufacturer can. Thus, policymakers must rethink liability concerns with an eye toward safety and societal benefit.

Another interesting consideration is the social interface between autonomous cars and cars with human drivers. Particular robotic driving styles might cater better to human acceptance and welcome. Would passengers be upset if their cars insisted on following the posted speed limits instead of driving at the prevailing speed of traffic? If we are sharing the road, would we want them to be servile, always giving right of way to human drivers? These are considerations we can evaluate directly using the tools of Social Robotics, evaluating systems with varied behavioral characteristics using real humans as study samples. Sometimes these studies have non-intuitive results. We might find that what was meant to be polite hesitance might be interpreted as lack of confidence and might cause other drivers to question the autonomous car's capabilities, feeling less safe in their vicinity. Traffic enforcement for autonomous cars could be harsher in order to send a message to the manufacturers, or more lenient because officers assume transgressions were

the result of computational errors, rather than intentional violations.

As the examples above highlights, when we put robots in shared human environments, social attributions become relevant to the robot's welcome and effectiveness because of what they communicate and the reactions they evoke. Pedestrians frequently make eye-contact with drivers before crossing a road. An autonomous car should be able to signal its awareness too, whether by flashing its lights or with some additional interface for social communications.

A robot should also be able to communicate with recognizable motion patterns. If a driver cannot understand that a robotic vehicle wants to pass her on the highway, she might not shift lanes. If the autonomous vehicle rides too close behind someone, that person might get angry and try to block its passage by traveling alongside a vehicle in the next lane. Such driver behavior in isolation might seem irrational or overly emotional, but it actually reflects known social rules and frameworks that machines will need to at least approximate before they can hope to successfully share our roads.

"Policymakers can make better choices to ease human anxiety and facilitate greater acceptance of robots."

Conclusion

This next generation of robotics will benefit from pro-active policymaking and informed, ethical design. Human partnership with robots is the best of both worlds—deep access to information and mechanical capability, as well as higher-level systems thinking, ability to deal with novel or unexpected phenomena, and knowledge of human salience. Ultimately there will be no single set of rules for collaborative robots. The need for customization and individualized behavior systems will occur both because it is what we want (consumer demand), and because it works better. New applications and innovations will continue to appear.

The unique capacities of humans and machines complement each other, as we have already seen with the proliferation of mobile devices. In extending this design potential to socially intelligent, embodied machines, good design and public policy can support this symbiotic partnership by specifically valuing human capability, maintaining consideration of societal goals and positive human impact. From considerations of local customs to natural human responses to agent-like actions, there are deep cultural considerations impacting the acceptance and effectiveness of human-robot teams. As social robotics researchers increase their understanding of the human cultural response to robots, we help reveal cultural red lines that designers in the first instance, and policymakers down the road, will need to take into account.

In better understanding the design considerations of collaborative robots, policymakers can make better choices to ease human anxiety and facilitate greater acceptance of robots. Regulations encouraging and safeguarding good design will influence whether users will want to keep humans at the helm, as in telepresence systems, join forces in a shared collaborative environments, or hand over the decision making while riding in an autonomous vehicle. Clear interfaces and readable behaviors can help the human partners better understand or direct the robot actions. Improved and reliable perception systems will help machines to have reliable local autonomy and interpret the natural behaviors and reactions of their human partners. The goal is to create guidelines that allow social and collaborative robotics to flourish, safely exploring their potential for positively impacting our lives.

1. F. Castelli, et al. Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns. NeuroImage, Vol 12: 314-325, 2000.

2. A. Engel, et al. How moving objects become animated: The human mirror neuron system assimilates non-biological movement patterns. Social Neuroscience, 3:3-4, 2008.

3. F. Abella, F. Happéb and U. Fritha. Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, Vol. 15:1, pp. 1–16, January–March 2000.

4. C. Kidd and C. Breazeal. A Robotic Weight Loss Coach. Twenty-Second Conference on Artificial Intelligence, 2007.

5. J. Carpenter. (2013). Just Doesn't Look Right: Exploring the impact of humanoid robot integration into Explosive Ordnance Disposal teams. In R. Luppicini (Ed.), Handbook of Research on Technoself: Identity in a Technological Society (pp. 609-636). Hershey, PA: Information Science Publishing. doi:10.4018/978-1-4666-2211-1.

6. M. Mori. Bukimi no tani. Energy, 7(4), 33–35. 1970. (Originally in Japanese)

7. N. Kitano. Animism, Rinri, Modernization: the Base of Japanese Robotics. Workshop on RoboEthics at the IEEE International Conference on Robotics and Automation, 2007.

8. C. Mims. Why the Japanese love robots and Americans fear them. Technology Review, October 12, 2010.

9. B. Robins, et al. Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction

skills?. Universal Access in the Information Society 4.2 (2005): 105-120.

10. K. Darling. Extending Legal Rights to Social Robots. We Robot Conference, University of Miami, April 2012.

11. A. Prakash & W. A. Rogers. Younger and older adults' attitudes toward robot faces: Effects of task and humanoid appearance. Proceedings of the Human Factors and Ergonomics Society, 2013.

12. "See a drone? Shoot it down." Washington Times, December 10, 2013. http://www.washingtontimes.com/news/2013/dec/10/see-drone-shoot-it-down-says-colorado-ordinance/

13. R. R. Murphy and J. L. Burke. The Safe Human-Robot Ratio. Chapter 3 in Human-Robot Interactions in Future Military Operations, F. J. M. Barnes, Ed., ed: Ashgate, 2010, pp. 31-49.

14. J. M. Beer & L. Takayama. Mobile remote presence systems for older adults: Acceptance, benefits, and concerns. Proceedings of Human-Robot Interaction Conference, 2001.

15. C. Breazeal, A. Brooks, J. Gray, M. Hancher, C. Kidd, J. McBean, W.D. Stiehl, and J. Strickon. "Interactive Robot Theatre." In Proceedings of International Conference on Intelligent Robots and Systems, 2003.

16. H. Knight, V. Ramakrishna, S. Satkin. A Saavy Robot Standup Comic: Realtime Learning through Audience Tracking. Workshop paper. International Conference on Tangible and Embedded Interaction, 2010.

17. B. Mutlu and J. Forlizzi. Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. Proceedings International Conference on Human-Robot Interaction, 2008.