The Tactile Ethics of Soft Robotics: Designing Wisely for Human–Robot Interaction

Thomas Arnold and Matthias Scheutz
Human-Robot Interaction Laboratory
Tufts University

ABSTRACT

Soft robots promise an exciting design trajectory in the field of robotics and human–robot interaction (HRI), promising more adaptive, resilient movement within environments as well as a safer, more sensitive interface for the objects or agents the robot encounters. In particular, tactile HRI is a critical dimension for designers to consider, especially given the onrush of assistive and companion robots into our society. In this article, we propose to surface an important set of ethical challenges for the field of soft robotics to meet. Tactile HRI strongly suggests that soft-bodied robots balance tactile engagement against emotional manipulation, model intimacy on the bonding with a tool not with a person, and deflect users from personally and socially destructive behavior the soft bodies and surfaces could normally entice.

Introduction

Soft robots promise an exciting design trajectory in the field of robotics and human–robot interaction (HRI), promising more adaptive, resilient movement within environments as well as a safer, more sensitive interface for the objects or agents the robot encounters. Although safety concerns have typically led robot designers to develop sophisticated hard body designs and control strategies to minimize the risk of physical harm to the robot and person alike, accidents from hard-bodied robots do occur. Soft-bodied robots ostensibly can pose less of a hazard in these socially interactive contexts, both to immediate touch and to deformability of the robot as a whole (similar to an automobile knautschzone). In terms of touching and being touched by objects in the environment, of course, it is human beings who represent the most important horizon for soft robotics' advances. Tactile HRI is a critical dimension for designers to consider, especially given the onrush of assistive and companion robots into our society. The more one considers the many functions and purposes to which socially engaged soft-bodied robots might be put, the more it appears that the usual measures of safety and durability are just the tip of the iceberg when it comes to responsible design. In this article, we propose to surface an important set of ethical challenges for the field of soft robotics to meet. Tactile HRI strongly suggests that soft-bodied robots balance tactile engagement against emotional manipulation, model intimacy the bonding with a tool not with a person, and deflect users from personally and socially destructive behavior the soft bodies and surfaces could normally entice.
Soft Robotics Within HRI Ethics

Soft robots present opportunities for novel locomotion and manipulation across many application domains. The goals of more flexible, adaptive, and resilient action can, when the robot function entails human interaction, align with perceptual goals, that is, achieving heightened sensory intake. To those ends, the technical challenges of making artificial surfaces more life-like, if not human-like have commanded considerable research attention. Mimicking human motions and receiving more fine-grained tactile information seem to be a technical package. As one projects some of these goals into concrete social contexts, however, one can recognize that there may be competing purposes or interests at work.

Softness and usability have most notable featured in HRI research as a therapeutically geared venture. Paro is perhaps the most notable of products in this vein, whether for elder companionship, children with autism, or those suffering from post-traumatic stress disorder. The combination of its soft, plush surface with additional encouraging feedback (small vibrations or indications of calmness) creates a soothing presence that researchers and care providers have been keen to explore for possible benefits. The Huggable is another case of establishing touch as a key to human–robot relationship, one situated within Breazeal's stream of companion research. This research has produced a number of prototypes that are especially engaging to children, and these promise (perhaps in a softer form of Jibo, Breazeal's entry into the domestic robot market) to enter into more settings where general companionship is the objective.

Even within that range of application, then, one may ask how the technical achievements of soft robotics are to be measured. Is it therapeutic effectiveness as measured by caregivers, or personal feedback from the users themselves? HRI research has started to broach the issue of tactile interaction and the ways robotic design can best suit it (e.g., how one would convey emotions to a robot through gestures of touch). Still, the potential conflicts involved have not surfaced with much deliberate treatment. Whether outlining the different kinds of touch that can occur or examining particular forms for a soft robotic technology like artificial skin, the closest thing to a general objective for this research would be the laudable goal of providing therapists with an increased repertoire of tools and devices to assist their work. But there are many nontherapeutic ways in which soft robots, because of the interactions they may foster, can affect social contexts of HRI, and the ethical implications of that possibility are worth spelling out in more detail.

Primacy of Touch

Although HRI research has recognized the Uncanny Valley for some time, it is worth noting how oriented toward visual and aural, not tactile, aspects of a robot the scholarship on that valley is. A turn to the tactile is more than due, given how primal and central touch is to human development. Like all higher mammals, human beings have strong intrinsic tendencies toward building basic communicative functions and attachment relationships through touch. Touch is a two-way channel through which caregivers and those cared for orient to one another, and more sophisticated gestures of grooming, affection, and conflict build upon those foundations. It is also clear that socially significant touch is not always between people: a child clutching a stuffed
animal can be a key embodiment about his or her relationship with the world, just as aggressively squeezing it may be. Recent research shows how much even holding a warm cup can translate into more positive social feelings toward the person with whom one is conversing, one among several examples of haptic experience having more influence than we would imagine.

Touch is an ongoing, if not always conscious, means to organize and direct our social experience. Even in the sphere of popular culture, the form and softness of robots carry strong messages about attachment and vulnerability. The box office hit *Big Hero 6* (whose producers consulted with soft robotics researchers) centers on a robot whose inflatable, puffy body conveys both cuteness and adaptability (and whose heroic actions often rely on the power of resilient softness). A fascinating article by Aragon *et al.* demonstrates how cute and cuddly objects can excite decidedly aggressive responses. If cute and engaging robots are made of soft materials, are they liable to be squeezed too much (both for their own functioning and, more importantly, the well-being and mindset of the user)?

As HRI scholarship moves forward in investigating how these tactile, embodied features affect a robot's performance with people, it would do well to distinguish and integrate three layers of how robots present themselves interactively. The *appearance* of a robot is perhaps the most quickly apprehended, and HRI scholarship has tested many different kinds of form, bodies, faces, facial expression, and basic physical gestures by way of seeing which ones are comparatively likeable, engaging, functional, and so on. The next layer one could just call *behavior*, encompassing its forms of movement, verbal communication, and other social actions (turn-taking, offering help, etc.). At a final level, one that HRI must investigate more thoroughly is the *experienced behavior or disposition of the robot*: How a robot's whole presence, including its physical presence, affect how people view it is key for thinking through what features and qualities are the most important to seek. Material and forms whose softness draws more aggressive, reckless, or addictive touch (or some combination) will alter what demands are being put on the robot's programming and overall ability to succeed in charged contexts.

The nascent developments in soft robotics lay the ground for us to incorporate ethical reflection on what purposes texture, firmness, and function should serve, given how users and others are affected by the technology. To establish a framework for the social and ethical implications of soft robotics for HRI, we discuss the following three rubrics: (1) bonding, (2) specifying function, and (3) modeling in relation to social norms.

**Bonding and Attribution**

A growing body of scholarship in HRI has highlighted how easily people can attribute emotionally charged personal qualities to a robot, even when it is fairly clear that the robot cannot reciprocate feelings of any sort. It need not be humanoid shape or voice. The “unidirectional” bond is a phenomenon that continued to draw scrutiny, and some commentators question how much of a danger it represents. Although it may seem extreme that soldiers hold funeral for IED detectors, or give gifts to vacuum robots, we can form many types of more or
less intimate relationships with inanimate objects. Not only do children have favorite dolls and toys to cuddle and talk to, but even adults have favorite mugs, tools, clothes. Some name their cars. So how much difference would it make for a robot to be softer in surface and/or structure? Given recent research into how touching robots, consciously or not, can stimulate human subjects, it is even more pertinent to investigate how feel and texture might heighten or dampen basic attributions toward a robot.

A softer feel in and of itself may be pleasing or comforting to a person interacting with a robot, and may elicit a response of trust and openness. This effect may be compounded with certain forms and compelling voice ability of the robot, catalyzing touch into an even stronger conduit for bonding. The critical point for soft robotics research for HRI to explore is what lifelike or desired softness actually cultivates in terms of a relationship, and how softness works in tandem with other robotic features. The feel of softness can be the result of a long-term interaction that has settled into deep familiarity and comfort through use, without there being any projection whatsoever of emotional reciprocity. Indeed, such closeness might be a healthy model for a robot: a tool whose use is intimately understood and whose usefulness is more fully tapped accordingly.

Although much of what soft robotics centers on is tactile interaction, it is worth adding that a soft, pliable form may compound tactile-based attribution through the sight of its movement. Maneuvering around or pushing off objects, moving with more subtlety and grace (whether humanoid twisting and gliding or animal-like slithering) may elicit attributions of vitality, resilience, adaptability, and power. These could combine with the tactile experience a person might have of the body to make for more complex, relational feelings of intimacy or estrangement. Even the locomotive aspects of soft robots, then, should take into account how people interacting with them will, for better and worse, project agency and patience toward them. Tactile-based bonding facilitated by soft robots will connect to fuller depths of embodied social interaction, even subtleties of proxemics and temperature.

**Functional Orientation**

The issue of bonding leads naturally into the larger question of what social robots are meant to accomplish. Given the kinds of attachment that touch might engender, an ethical survey of how soft robots should best serve tactile ends requires more perspective on the function of the robot as a whole. A therapeutic robot's function is presumably tied to various forms of rehabilitation, the improvement of mental or emotional health. That can lead to more confined assessments of how, for example, tactile interactions should take place for children with autism. There may be direct specifications tied to therapeutic objectives, with empirical testing of how well various textures, feels, and responsiveness help achieve those ends.

Social robotics is fast spilling over such identifiable lines, whether by market forces or by other motivations. As we discuss in more detail hereunder, amid questions of whether to promote or ban sex robot technology, we found in a recent survey that views on sex robots showed interesting divides as to the purpose of such robots and their impact on relationships. To the degree such robots are not viewed as replacements, employing soft robots to replicate a human
body's feel may not be functional. Recent research into tactile response and shifting touch preferences even suggests that our cognitive mindset around function might influence what we like to touch. For soft robots, as with other design features for social robots, this should give pause to the common talking point (and sci-fi trope) that the peak of robotic advancement is them being indistinguishable from human beings.

Even for less charged robotic technology than sexbots, function will be a key axis for directing and assessing soft robot designs for touch. Long-distance communication through The Hug presaged a range of current efforts to send physical affection through a robotic conduit. It would seem like replication would be the main function of such technology, such that one person feels like the other person is hugging them. But it is also true that the purpose is to communicate a person's hug, not to replace the need for that person altogether (or have the recipient prefer The Hug to the sender in person)—when does replication divert the ostensible purpose of the robot?

These questions of function will be critical to work through and apply to the use of soft material technology for social robotics. As social robots are purchased privately, especially for domestic use, the lack of control over function may lead to behavior that would seem to violate a robot's legitimate purpose; for example, the robot may be an effigy for severe physical abuse and violent role play. Sex robots may invite that use, some have pointed out, just as telecommunicative touch might exacerbate a trend toward mediated interaction through devices instead of keeping in-person company. As domestic companion robots roll out of production and into homes, it will be all the more pressing to ask what their proper role is once they get there. Soft robotics should be as responsive to society's deliberation on social robots' proper function, as heated and complex as those discussions may be, as any other facet of robotics.

Habits and Modeling

The possible elusiveness of a single function, and the prospect for dysfunctional behavior with the robot on the part of users, leads us to consider how soft robotics might enable not just an individual action but longer lasting habits and orientation toward robots. Ethically speaking, this question leans on the idea of modeling behavior, wherein forms of action start to affect the social sphere. It speaks of how children, for example, might pick up a way to treat a robot from another, or how a soft robot in the public sphere might be touched by passersby or users.

One dynamic that could result from a soft, inviting body is the enticement for physical testing, causing the robot to undergo stress to see how it will react. Does the soft quality of its body and surface invite more of that treatment, both because (1) it does not hurt the interactant with a hard or rough surface and because (2) it promises to withstand the treatment without serious damage. A comparison could be made with bubble wrap—children may not be hurt by bubble wrap on the floor, but they could quickly create a noisy, distracting classroom, all while fighting among each other about who gets to pop more.

In this case of a social robot, certain features of softness could mesh with models of interpersonal behavior, or perhaps behavior with an animal. As already discussed, there may be models of attachment that are not true to the robot's actual abilities and processing, e.g. caring for or
sympathizing with the robot. But what if the companion just becomes an object of abuse, a physical means to vent frustration? Moreover, what if the robot generates a certain kind of perverse aggression precisely by being soft and indestructible? These types of dynamics have been studied more by child psychologists, and it is worth connecting with psychological literature to ask how such dysfunctional and antisocial behavior might feature in the soft social robot's operation among people.

One opposing point for why soft robots might be more preferable than hard-bodied robots is that softness might indicate more vulnerability. Squeezing or cuddling a soft robot might seem more inviting, but rougher treatment might befall a robot whose hard surface suggests impermeability and invulnerability. To be sure, there are different kinds of abuse that a hard-surfaced robot could lead a user to inflict if frustrated or curious. Nonetheless, it is crucial to balance any associations of safety and harmlessness that softness carries with the less productive reactions a soft robot could encourage.

**The Case of Sex Robots**

The main point of this article is to show that ethical challenges will pervade the development and implementation of soft robotics in robots. This discussion spans all social application domains, not just extreme or special cases of robotic action. That said, one way to drive the ethical challenge home is through a particularly powerful illustration, in this case the headline-grabbing case of sex robots. Our recent research on people's opinions of sex robots lends empirical substance to what might seem too hypothetical a set of ethical concerns. But as our opinion survey results suggest, sex robots point toward many issues of tactile intimacy, physical presence, and social values that could easily apply to tactile interaction writ large. It is worth zooming into these results to see how connected many kinds of social robots will be to dynamics people expect sex robots to involve.

Our surveys were conducted amid increasing attention to the potential risks and opportunities sex robots might represent for society. The *Campaign Against Sex Robots* has taken a firm stand against the dehumanizing effects sex robot technology could have on actual people. Countering voices have claimed sex robots could serve people otherwise excluded from sexual activity, to the point where sex robots might replace humans in the bedroom within decades. What was missing were systematic looks into what opinions the public has about sex robots: what they are, what appropriate uses they could have, and how interactions with them are to be categorized and judged.

Our first iteration of the survey found significant gender effects in terms of how appropriate various uses of sex robots were. There was general agreement, however, about what abilities sex robots would have (e.g., language for dialogue) and which robot forms would be appropriate (e.g., child forms were agreed upon as inappropriate). There were also particular contexts wherein women and men were closer in their judgment that sex robots could be appropriately used: (1) assisting a relationship between human beings (e.g., spouses), (2) training (e.g., for sex harassment prevention), and (3) contexts of extreme isolation (e.g., space vessel, base station).
The relationship context and societal aim involved, in other words, were part of how sexual interaction with a robot was judged.

The results of two subsequent studies have supported the idea that the use of sex robots is not cordoned off as an impersonal act—sex is related to many kinds of touch and intimacy, and indeed the uses of sex robots fall under moral arguments for helping others, maintaining trust, sustaining one's needs in addition on one's actual romantic partner (i.e., a human being), and for not modeling bad behavior.\footnote{33} The second study added questions about the advantages and disadvantages of sex robots. The subjects' responses strongly suggested that the issues of bonding, function, and modeling are in how sex robots are appraised morally. Although sex robots themselves were not thought of as real threats to replace human–human sex altogether, nor in and of themselves a risk for exploitative bonding, their advantages and disadvantages were tied to how they could threaten or enhance human relationships. Moreover, these impacts were gauged not just by the individual but also how larger social patterns of behavior could be affected. We again found that some uses and forms were ruled out of bounds (child shapes), but even when generally regarded as inappropriate, women allowed for certain social goods to be pursued through their use (training, extreme isolation). Our third study makes an initial venture into cultural comparison (with more research still being called for), comparing U.S. with European subjects on the second survey’s questions. We found some broad agreements between both sets of respondents around the appropriateness of sex robots, their assumed qualities and abilities, and the contexts and conditions wherein they are acceptably employed. Beyond the U.S. context, then, sex robots elicit ethical reflection that takes larger interpersonal factors and social contexts into account.

When it comes to the contribution that soft robotics might make to sex robots, then, we are not to think of mere custom-built oddities that only a user will experience in the bedroom. Sex robots could well have larger circles of impact through the tactile interactions they cultivate, whether within a relationship or in wider spheres of society. Part of the social modeling for robots that needs to be considered is the way that ostensibly “non-sexual” roles can be social eroticized or sexualized. Because of how sexual interaction is woven into networks of social relationships, it is naive to think that socially interactive robots will not be open to sexual construal and use. Aldebaran, the maker of the robot Pepper, has already recognized this dynamic: they require users of Pepper to sign a contract that includes a prohibition against sexual use.\footnote{34}

As we pan back to social robots, then, the ethical upshot of physical interactions with sex robots must extend to how social robots, however designated in terms of function, might actually be used. Gauging its performance unthinkingly by how lifelike its features are is is not sufficient for wise design—one must consider that tactile interaction will bear not only on the individual user's tactile experience but also what larger societal ends and relationship dynamics that touch could affect. The touch between a person and a robot, in other words, carries with it the implicit connection to human–human or other forms of touch—how that person will want to touch and be touched in the rest of his or her daily life, and how his or her touching and being touched features for better or worse within a community at large.
Guidelines for Soft Robot Design Addressing Social Interaction

On the basis of the preceding considerations, we propose three general guidelines for situating soft robotic technology within many domains of social HRI. Soft robotics should look toward (1) developmentally oriented attachment (with information gathering kept separate from patiency), (2) primary fidelity to function, and (3) appropriately bounded social modeling. These rubrics can inform approaches to soft robotics to be more useful, humanizing, and exemplary of societal norms.

In light of the many ways social robots might serve, including a capacity of being companions, there is a spectrum of attributions they will invite. The therapeutic and developmental roles they may fill mean that attachment or bonding is not itself to be avoided at all costs. The question is, then, what parameters or qualities should govern how such attachment takes shape. These should represent both the risks and opportunities of tactile attachment, safeguarding against destructive or maladaptive projections while encouraging life-enhancing experiences and abilities for those with whom the robot interacts.

That said, a better approach to the ethics of soft robotics than risks versus opportunities is that between touch as perception and touch (and, to a lesser extent, motion) as an action. In the former, one can consider touch as a means of gaining information about the environment and interactants. The design imperative, accordingly, is to ensure that soft robotic technology—in skin, body firmness, and other means of sensing tactile input—is experienced as gathering important information. Whatever attachment or familiarity might develop between a person and a robot, it cannot stem from soft robotics suggesting animal-like, not to mention person-like, patiency. Feeling the touch of others is a robotic conduit for the larger purposes of the system's designers and implementers (therapeutic, companionship, education, etc.). There should be no suggestion, however implicit, that the robot suffers or enjoys the tactile feedback. Although actual physical damage may be a fact that the user learns (perhaps even from the system itself), no analogy to pain or discomfort should be encouraged by soft materials and their responsiveness. As a result, there may need to be accompanying reminders or cues in the tactile interaction itself (whether lights, colors, sounds, or movements) that accompany soft robots' employment and establish its incongruence with typical patiency.

In a similar way, any initiated touch or interactive movement (which would be enabled by being soft bodied, whether flexible locomotion or maneuvering around objects) on the part of the robot should convey a fidelity to its overall function in the social context. The touching of a patient, say, or maneuvering among people in a public space, can and should communicate larger goals or purpose, especially for the sake of coordinated action. At the same time, the system as a whole should not exploit its functional features—soft robotic technology no less than any other aspect—to represent itself as having its own desires or competing interests as an autonomous system. Manipulative or exploitative relationships could form if people are induced to view a robot as having its own inherent desires, or rights, or even dignity, which could depend, in part, on how well its touch mimics living organisms (not just human beings but also, e.g., tenderness from pets). Just as a robot's response to touch should convey processing of physical force, but not dramatize that as having felt it like a human being would, so soft robotics should strive for imitative touch and movement gestures only after grasping more fully their relational consequences. As with perceiving touch, administering touch may need means of disfluency and
attribution–interruption to keep the physical virtues of touch from bleeding into inappropriate projections toward a robot of having a will or interests alien to the human beings employing the robot.

Heading off attributions of independent will or purpose is obviously critical for therapeutic uses. For the client or patient, especially as a child, having a safe and reliable environment of treatment means preventing uncanny or oddly alienating actions or qualities on the part of the robot. Even at a tactile level, having touching gestures veers off from a process that is clearly understood by caregivers, and support networks will impede the full benefits that a robot will provide. For other contexts, for example, domestic companionship, the task will also demand better articulation of what the robot's role really is and should be. It is worth pointing out that the domestic companion market has not provided anything approaching that level of detail to assist soft robotics engineers going forward.

Recognizing that certain social actions can extend beyond the confines of a particular one-on-one interaction—either through onlookers or other interactants picking up on a behavior or the person carrying an interaction into other contexts—we would urge that social modeling be accounted for as a feature of soft robotic technology in the social sphere. Situating soft robotics within a robotic system is not enough, because we must also situate the robot in a social context that fills out how its interactions should unfold. The inducements to aggressive or gratuitously risky behavior on the part of soft robotics must not impede or displace constructive patterns (say, reporting health symptoms in an eldercare home) that could form positive social practices. More work in HRI will obviously be needed to determine where the real differences are in how soft robotic technology shapes interactions, but one can never lose sight of how its affordances and usages both reflect and reshape the societies using them.

For all three of these guidelines, it may be tempting to point toward software as an “out” from tactile quandaries. Perhaps a robot would have sufficient natural language processing and speech generation capabilities, in addition to planning and goal managing features, to dissuade people from dysfunctional forms of interaction. Given the primacy of touch discussed earlier, however, this is a dubious tack to take. It is more likely that touch will be more influential than users even realize. The material quality of the robot may overpower the interactive models afforded by the software: this is the daunting prospect that the design of social robots must confront and engage. Being careful with that material, including the soft robotics technology that a robot might possess, is an ethical imperative.

Conclusion

The advancement of soft robotics will unavoidably go hand to hand with greater physical expectations for socially interactive robots. For more intimate and sensitive interactions, social robotics will draw upon that advancement, and the ethical challenge will be making sure that usage will stay wedded to genuinely constructive purposes and broader societal interests. The dynamics we have explored here collectively point toward an unavoidable, indeed sobering, ethical upshot: the material quality of a robot may fundamentally alter the efficacy of its programming. In other words, a hard-bodied robot and soft-bodied robot may share the exact
same code for dialogue, movement, gestures, and execution of physical tasks (lifting, cleaning, etc.). But with the soft-bodied robot, if it elicits closer bodily presence and physical contact (e.g., affectionate or aggressive squeezing), performance may be compromised or perhaps taxed with an additional set of social and physical burdens (having a person in its space of operation, not being to move freely while embraced, etc.). This must be stressed: soft robotics is not just a rough translation of actions defined by machine code but also an interface that could change what is expected and what is practically possible for that code to express in robotic action.

The technical challenges of developing more engaging and lifelike surfaces and forms for robots are, then, not the exclusive or ultimate measure of soft robotics. For better or worse, such challenges must be evaluated relative to how the quality of HRI they promote will affect human lives. Tacit assumptions about the purpose for which soft robots will be used should be brought into the light of ethical examination, so that we do not simply build human-like skin but do so in light of how and why a “skin” or “body” can create more effective, helpful interactions. Far from discouraging soft robotics research or downplaying the importance of its results, the ethical challenges of its use for social robots should enrich considerations of how soft robotic technology can best serve.

NOTES