Ethical Principles for Socially Assistive Robotics
David Feil-Seifer and Maja J Matarić

**Interaction Laboratory** 

Center for Robotics and Embedded Systems

Department of Computer Science, University of Southern California

Los Angeles, CA USA 90089-0781

dfseifer@usc.edu —mataric@usc.edu

#### **Abstract**

The health, education, and other service applications for robots that assist through primarily social rather than physical interaction are rapidly growing, and so is the research into such technologies. Socially assistive robotics (SAR) aims to address critical areas and gaps in care by automating supervision, coaching, motivation, and companionship aspects of one-on-one interactions with individuals from various large and growing populations, including stroke survivors, the elderly and individuals with dementia, children with autism spectrum disorders, among many others. In this way, roboticists hope to improve the standard of care for large user groups. Naturally, SAR systems pose several ethical challenges regarding their design, implementation, and deployment. This paper examines the ethical challenges of socially assistive robotics from three points of view (user, caregiver, peer) using core principles from medical ethics (autonomy, beneficence, non-maleficence, justice) to determine how intended and unintended effects of a SAR can impact the delivery of care.

#### Introduction

The most obvious and direct risk of any assistive technology, including socially

assistive robots, is the potential of physical harm. While this is an important risk to examine, SAR is primarily concerned with robots that provide assistance through social, rather than physical interaction. In this paper, we outline commonly accepted core principles from medical ethics and use those principles as guidelines for evaluating the risks of SAR. We use examples of SAR systems to describe the ways that robots are currently being used as directions for future use based on ongoing research. We then discuss the core ethical principles to be examined. Finally, we apply each principle to SAR in turn and discuss its implications.

## Definition of Socially Assistive Robotics

Socially Assistive Robotics (D. Feil-Seifer & Matarić, 2005, 5) describes a class of robots that is the intersection of Assistive Robotics (robots that provide assistance to a user) and Socially Interactive Robotics (robots that communicate with a user through social, non-physical interaction). Assistive robotics is a broad class of robots whose function is to provide assistance to users, ranging from getting out of bed, brushing teeth, locomotion, and rehabilitation. Socially Assistive Robotics (SAR) is the class of robots that provide assistance to users primarily through social, rather than physical, interaction. This section provides examples of SAR systems.

Wada et al. (2003, 23) describes the design of Paro, a robot for pet therapy applications for nursing homes that do not allow pets. Pet therapy has been shown to have a positive effect on the elderly in nursing home settings (Perelle & Granville, 1993, 16) but there are logistical challenges to having animals into nursing homes. Paro was built to resemble a baby harp seal, and designed to interact like a pet, with simple sounds and movements made in response to being held and petted. Experimental results suggest that Paro may be effective at reducing stress in nursing home residents. In addition, when placed in common areas of nursing homes, it produced increased social activity among residents. This suggests that SAR systems may be useful not just for their direct

therapeutic applications, but also as catalysts for social interaction more generally.

Another SAR system is Roball (Salter et al., 2007, 18), a self-propelling robotic ball that can sense its position and motion and thus the way it is being played with. Roball is being evaluated for use by children, including children with autism spectrum disorders (ASD) in the home or in clinical settings. Children with ASD typically have decreased social interactive behavior; encouraging play with therapists, family, and peers could have both diagnostic and therapeutic uses. Roball and other robots for play could be used as an addition to current ASD diagnostics or therapeutic regimens, or as tools for developing new diagnostic and therapy methods. In general, the aim of SAR for ASD is to encourage children to initiate and sustain social interaction (Robins et al., 2007, 17) with a parent, therapist, sibling, or peer.

Rehabilitation post stroke is another domain where socially assistive robots can provide therapeutic benefit. Rehabilitation robotics has been developing robot arms that apply and measure forces on the user's limbs. Such hands-on movement training is particularly useful in the early stages post stroke. However, a major long-term challenge of post-stroke recovery, and rehabilitation in general, is encouraging compliance with the prescribed therapeutic regimen. Matarić, et al. (2007, 12) describes a SAR system designed to improve therapeutic compliance through verbal, non-contact coaching and encouragement. Such systems are designed to work in concert with established stroke exercise methods, such as Constraint Induced Therapy, building on and augmenting effective health care practices.

Concurrently with developing SAR technologies, ethical appraisal studies are being conducted about their acceptance. Mutlu & Forlizzi (2008, 14) conducted an ethnographic study of a delivery robot used in multiple departments of a hospital, finding that different patient groups had very different reactions to the robot. For example, cancer units were not accepting of the robot, finding it annoying, while postpartum units were accepting of the robot, calling it delightful. The results of this study suggest that user populations could have completely different experiences with the

same robot, and that these experiences could be based on the users' pre-existing social and task dynamics and context. Tapus et al. (2009, 21) described a study in which elderly participants with Alzheimer's Disease interacted with a SAR robot that promoted cognitive exercises through a song recognition game in a 6-month study. The study participants included the robots in their narratives and preferred it to a computer. Turkle (2005, 22) demonstrated that some participants interacting with robots can correctly identify the robot's intended emotional abilities and operational capabilities. These participants could also correctly distinguish equivalent capabilities in a person, pet, or other relational artifact. However, it was also demonstrated that some users formed attachments and emotional bonds with robots they were interacting with. These attachments led to misconceptions about the robots' emotional capabilities. For example, one user felt that the robot would miss him when he was gone, something that the robot was not capable of doing. In their hyperbolic, yet poignant, article, Sharkey & Sharkey (2010, 19) argue that such attachments in children could lead to malformed development and emotional problems.

## Persons affected by SAR

Socially assistive robots are being designed for use in a wide variety of settings, including hospitals, schools, elder care facilities, and private homes. The intended endusers of such systems are individuals with special needs, but SAR systems must operate in real-world environments that may also include family, caregivers, and medical personnel. Consequently, the effects of SAR must be assessed for all of the individuals affected by the technology.

# Core Ethical Principles

There are many ways to approach potential ethical issues related to technology in

general and SAR in particular. Several appraisals of specific SAR systems have been implemented and some have discussed ethical dilemmas that a particular system poses (Sharkey & Sharkey, 2010, 19; Turkle, 2005, 22). Studies have also aimed to establish ethical benchmarks related to the design, manufacture, or use of SAR (Feil-Seifer et al., 2007, 7; Kahn et al., 2006, 9). Finally, some appraisals have applied core ethical principles to identify potential problems (Arkin, 2008, 3). In this work, we apply an established medical ethics framework to identify potential issues related to SAR. This framework uses the following core principles for considering ethical issues (Beauchamp & Childress, 2001, 4):

- Beneficence caregivers should act in the best interest of the patient;
- Non-maleficence the doctrine, "first, do no harm," that caregivers should not harm a patient;
- Autonomy the capacity to make an informed, un-coerced decision about care;
   and
- Justice fair distribution of scarce health resources.

There is dissension about whether or not the Beauchamp and Childress model is the ideal model for assessing medical ethics, as the foundation for current ethical appraisal and ethical training, we feel it is a sufficient jumping-off point for discussion.

These principles underlie ethical reviews of experiments with human participants and can thus also provide broad categories for examining ethical issues related to socially assistive robotics. To perform such an examination here, we use examples from actual SAR system experiments. However, these descriptions are only considerations of hypothetical scenarios, and not meant to make judgments on the ethical validity of those specific SAR systems. In the next section, we describe the principles of Beneficence and Non-Maleficence and how they relate to the ethical use of SAR.

### **Beneficence and Non-Maleficence**

The principles of beneficence and non-maleficence state that caregivers should act in the best interests of the patient and should do nothing rather than take any action that may harm a patient. These principles establish that the potential benefits of an ethical treatment should exceed the risks. SAR, like any technology, features some risks along with compelling potential benefits,.

As noted earlier, SAR technologies are typically non-contact, so physical risk, while usually the most obvious ethical concern, is not a major issue of concern. SAR systems are designed so the robot does not apply any forces on the user. The user, on the other hand, can touch the SAR system, and in some cases (as with Paro, see above), such contact is part of the therapy. However, in the majority of systems, no physical contact is involved, and the robot may not even be within reach of the user, though it is typically within social interactive space conducive to one-on-one interaction through speech, gesture, and body movement.

In this section, we examine some of the aspects of SAR technologies that are unique and ways in which SAR systems in particular might impact not only the user directly, but also others in the shared context. In particular, the most prominent non-physical risks posed by SAR systems include, but are not limited to, attachment to the robot, deception about the abilities of the robot, and influence on the amount of human-human interaction of a robot's user.

# Relationships, Authority and Attachment

It is safe to assume that a robot would not be the only caregiver/therapist for an assisted individual; typically, care is provided by human caregivers including professionals and family members. Thus, the SAR system impacts all of these individuals in various ways. For example, a robot that does something that a human

caregiver would otherwise do (e.g., providing encouragement for performing exercises), might have as much impact on the human caregiver as on the patient, through the reduction of tasks related to a patient or through the reduction of workplace monotony. Specifically, many SAR systems are being designed to reduce the burden and burnout of family members and other caregivers. A SAR system might also provide a benefit to a caregiver by monitoring multiple aspects of the patient and providing ongoing quantitative assessments.

Sharkey and Sharkey (2010) described another significant ethical dilemma that occurs when a user becomes emotionally attached to the robot. While establishing engagement and having the user enjoy interactions with the robot is a goal of SAR, attachment can also result in problems under certain circumstances. For example, if the robot's effectiveness wanes, its scheduled course of therapy concludes, of if it suffers from a hardware or software malfunction, it may be taken away from the user. The robot's absence may, in cases of attachment, cause user distress, and possibly result in a loss of therapeutic benefits. Attachment issues can happen with users of all ages, from children, to adults, to the elderly. Such issues can be particularly acute in users who cannot understand the causes for the robot's removal, but can arise even with users who have full understanding of the circumstances. Our experiments with SAR robots interacting with elderly users, and users with Alzheimer's Disease, mentioned earlier, demonstrate that such users do engage with robots and miss them when the robots are removed (Tapus, et al., 2009, 21).

## Perception and Personification of the Robot

As discussed above, one goal of an effective SAR system is to establish a relationship with the user that leads toward intended therapeutic goals. However, since the user cannot be fully informed about the limitations of the robot, the following issue

arises: "Is there deception inherent in the personification of a robot by a user or a caregiver?" Such personification could be unintentional, arising from the caregiver referring to the robot as him or her, ascribing feelings to the robot, and assigning the robot greater intelligence than it may have. Studies have shown that people quickly form mental models of robots they are presented with, much as they do of people. Those models are often incorrect, as they are based on what people know best: other people. The designers of the robot may purposefully manipulate the perceptions of the user toward therapeutic goals, or may not intend to do so at all; in any case, if such perceptions are incorrect, the user is deceived.

Deception is a risk created by the use of robots in assistive settings. Some roles of SAR systems are most closely associated with people, such as those of a therapist, companion, teacher, or coach. In those roles, the robot may be constructed to physically resemble and act like a human equivalent in those roles. In other scenarios, the robot may fill the role of a pet or toy, with physical form to match. While it may be assumed that the physical form of the robot is deliberately designed to evoke the desired type of relationship with the user, there can be unintended ways in which the robot is perceived and received by the user. Studies of the so-called "uncanny valley" already demonstrate that the level of human-like realism of the robot has unexpected impact on people (Kanda et al., 2004, 10). Similarly, the size of the robot has impact on the interaction and perceived role: studies have shown that robots that approach the height and size of the user are received with some trepidation compared to smaller embodiments (Lee et al., 2009, 11). The way the robot is dressed and accessorized can also influence how it is perceived; a robot in a lab coat and wearing a stethoscope might be perceived as being medically competent even if it is not.

The issues of physical appearance are in many ways just the tip of the iceberg; communication is also crucial. Whether the robot speaks, and if it does so with a synthetic or recorded voice, male or female, accented or not, and containing emotion or not, are all important parameters defining the nature of the interaction. These

communication parameters play key roles in how effective the robot will be in a SAR setting. In addition to speech and language, embodied expression consisting of gesture, body language, and facial expressions comprises another complex area of study in human-machine interaction. This myriad of SAR design parameters has important consequences on the role of the robot and the resulting human-robot interaction; there is much research to be done in defining how factors affect interaction in general and user care in particular.

The relationship between the user and the robot as defined by the role of the robot can lead to deception with regard to the robot, contributing to increased risks. For example, a user perceiving the robot as a doctor or nurse could lead to deception. Such deception could be harmful, especially considering that the robot's communication and decision-making abilities are not on par with a human caregiver. A user could also believe that a robot is capable of assisting him/her in ways that a human would when in fact it could not. For example, if a user perceives a robot as having the abilities of a doctor, that user could equate telling the robot a pertinent piece of medical information with communicating that information to the doctor, potentially resulting in lost information. Conversely, if the user does not perceive the robot as a knowledgeable authority, he or she may not accept suggestions or instructions from the robot, thereby subverting the therapy process and rendering the robot ineffective. To complicate the matter further, such loss of authority may not be instantaneous; the user may be amenable to working with the robot for some time, perhaps due to the robot's novelty to the user, but may later lose interest in the robot.

Another related aspect of user perception of the robot's abilities and authority is the issue of recognition and reporting of suspect behavior. Consider a situation wherein a user is in obvious distress. A human observer would, or should, know to report the situation to an authority capable of helping. A robot, however, may not have the ability to recognize alarming behavior, yet people around the robot may believe it does, and so may fail to act in response, assuming the robot would/could handle the situation. Before

any technology is deployed in an assistive setting, it is critical to establish to all involved what the capabilities of the technology are. However, as SAR and other technologies become more pervasive, uninformed bystanders will be exposed to them, and assumptions of full disclosure will quickly become unrealistic. In general, the issue of projected authority and role of the robot based on its appearance and behavior is complex, and one that could be the topic of study from a range of fields, including ethics, social science, and engineering.

## Changes to Human-Human Interaction

The work of Wada et al. (2003) demonstrates that SAR systems can result in increased amounts of human-human interaction. However, a robot could just as easily be an isolating factor (Sharkey & Sharkey, 2010, 19; Turkle, 2005, 22). Most current examples of SAR use a robot as an enhancement of the roles of current caregivers, not as their replacement, and as an addition to existing therapy, not its substitute. However, if the robot is used as a replacement or substitute for human care, then the robot might serve to reduce the amount of human-human contact. This is especially a concern if the robot is the only therapeutic influence in a user's life. For populations that are known to suffer from isolation, including the elderly or children with developmental disorders, robots might facilitate further isolation even while delivering a therapeutic benefit. We have argued that such use of technology as proxies for human attention is a real risk, but not one that is new or specific to robotics. Television watching and playing computer games are both poor substitutes for attentive parenting but neither the TV nor the games can be blamed. Similarly, ethical and productive use of SAR technologies will

necessarily put the burden on the caregivers to not abuse the technology.

#### Discussion

The core principles of beneficence and non-malfeasance are crucial for deciding if the use of a socially assistive robot is ethical and beneficial for a particular user. While there are many stated benefits for SAR in terms of encouraging social interaction, encouraging therapeutic compliance, providing a therapeutic intervention, and providing advice, there are potential ethical pitfalls. Properly describing the capabilities and the role of an assistive robot is critical for caregivers to assess the potential for harm. In addition, proper communication between caregivers and users of SAR is crucial to minimize unintended deception. Finally, when robots are first introduced to users, the possibilities for upgrades or modifications that would change the robot's appearance or behavior, and the fact that the robot might or will eventually be taken away, should be made clear to the user.

Generally speaking, deception should be minimized wherever possible in order to avoid harm to the user. But, as noted earlier, since human perception of any part of a robot (facial expression, voice, gesture, appearance, size, etc.) is not yet well understood, unintentional interpretation and possible deception are inevitable until our understanding of the human-machine interaction is thoroughly studied and characterized. It is thus critical to conduct detailed studies in realistic but monitored settings before commercializing these technologies in order to improve both the safety and the effectiveness of the designs.

Ethical ramifications of SAR are not limited to the balance between risks and benefits. SAR also poses challenges for the user's informed decision-making ability, as discussed in the next section.

### Autonomy

The core medical ethics principle of autonomy dictates that patients should be able to make informed decisions about their own care. Extending this principle to SAR, patients should be able to make informed decisions about socially assistive robots that are part of their care. As discussed in the previous section, several factors make it likely that a user may not be capable of being fully informed about the abilities and limitations of a particular SAR technology, and be aware of his or her own possibly biased perceptions of it. People might believe (or be made to believe) that the robot is more capable than it is, which can create barriers to making an informed decision about care. There are also valid concerns about a user's privacy with SAR, as with most other technologies. If a robot is not able to properly distinguish between confidential information (e.g., personal health information) and information that the user permits for release, then the robot may create an unintended violation of a user's privacy. In this section, we examine problems relating to informed consent and privacy that have ethical implications.

In order to provide the user with enough information to make an informed decision about a robot, a critical question is: "Are the capabilities of an assistive robot being correctly described?" If a description of how the robot will be used does not give the user the necessary information to make an informed decision about using the robot, then the caregiver is not behaving in an ethical manner. Consider the example of a companion robot for use in nursing home that does not allow pets. If the user is told that the robot is "just like a pet," but later discovers that in fact the robot only has a limited and small repertoire of behaviors, the user may become disappointed and feel lonely. However, this is not a simple issue; the robot vacuum cleaner, Roomba, is capable of very few actions related to floor vacuuming, yet studies have shown that users of the Roomba are attached to it and demand that it be fixed and returned when broken rather than that it be replaced with a new one (Sung et al., 2007, 20). Different users have

different expectations and so it is not necessarily possible to warn a user completely about his/her perceptions and bonding with the robot, positive or otherwise.

Similarly, the role of the robot and possible misconceptions about that role, described in the previous section, could lead a user to expect high-level human-like medical care from a robot. While the capabilities of the robot may be effective in a specific application domain, they are not comparable to a human doctor or nurse, who may be able to assist the user with decisions or consultations outside of the prescribed therapy. If a user is anticipating an inappropriate benefit for the cost of a robot that s/he is considering purchasing, then that user is not fully informed. The impact of the decision is even more important if the user is considering an application that uses a robot in place of, rather than in addition to, a human caregiver.

The authority of the robot is another sensitive issue for SAR. A robot's intended role as therapist may exert influence on the user, putting in question who is in control of the situation and interaction. The question of "Who is in charge?" must be addressed carefully, because the technology may require a level of authority in order to be effective. A user that is feeling stressed or is in pain must feel free to stop an exercise, for example, even if that is counter to the robot's advice. However, a SAR system's role in many contexts is to give direction to a user, requiring some measure of authority derived from expertise. A lack of balance between user autonomy and robot authority could create an ethical dilemma.

When discussing authority with respect to SAR, privacy is of utmost importance. A robot might not have sufficient capabilities to distinguish between privileged information and information that can be distributed. A robot may also lack the ability to distinguish between individuals who have the authority to receive information about the user, and those who do not. Patients seeking medical care have an expectation of privacy backed by legal protection. However, a robot might not be able to meet these privacy obligations. In particular, a user might not realize that a robot's camera could record video, or display video in another location, or that wireless transmission of video data

cannot be guaranteed to be completely private. People perceive a robot's cameras as having similar capabilities to human vision; this is a natural but false assumption. As discussed in the previous section, the robot might not know to communicate information that is critical to care or how to communicate privileged information in a discreet manner. Therefore, it is important to make sure that the capabilities of a robot are sufficiently explained so that a user has as well informed of a model of the robot's abilities as possible.

The use of SAR can also have a positive effect on a user's autonomy. An example from an assistive technology study describes how elders in independent-living situations were asked to allow cameras into their homes in order to allow for home monitoring for safety. The elders were uncomfortable with this process, as they did not want to be seen, especially in private places like the bathroom. The experimenters responded by using computer vision to monitor only the user's silhouettes (Anderson et al., 2006, 2), thereby providing sufficient information for the home monitoring task, but also allowing the users autonomy in choosing what information they wished to release. SAR could employ similar techniques for allowing users privacy, thereby increasing user autonomy.

#### Discussion

Preserving the autonomy of a person seeking care is a core ethical value. For the most part, the procedures for informed consent are sufficient for allowing a user's autonomy in decision-making regarding care. However, the potential for user deception can interfere with a user's informed consent. Currently, the appearance of a robot and its ability to sense its environment and communicate with others might not match. This mismatch might result in (unintentional) deception of the user as to the robot's capabilities, which, in turn, may affect the user's ability to give informed consent. To mitigate this, users should be presented with a clear description of the robot's capabilities as well as limitations, but they also must understand that their perceptions of

the robot, responses to it, and the attachments and relationships they form with it, are not fully predictable, just as they are not in human-human interactions.

#### **Justice**

The principle of justice governs the fair distribution of scarce resources. This can be a very difficult topic when discussing experimental treatments such as socially assistive robotics. The authors know of no SAR systems that are currently used outside of the research setting, so discussing actual cases in the field is premature. However, we can presume that, for the foreseeable future, robots will be somewhat expensive. Thus a question that should be asked is "Do the benefits of SAR outweigh the costs?" Like other proposed therapies, quality of life surveys, or other methods for assessing medical economy can be used to assess relative benefits, and costs can be weighed against improvements observed (Aaronson et al., 2004, 1; Wood-Dauphinee, 1999, 24). There does not seem to be a significant difference between calculating the costs/benefits of robots compared to other assistive devices.

Another justice-related issue when discussing robotics in socially assistive settings is the notion of responsibility: "Who is responsible when things go wrong?" While this might not traditionally pertain to the principle of justice, fair allocation of responsibility for SAR systems might be related to a fair allocation of therapeutic resources. When a robot does not behave as intended, it could be the result of user error, or it could be the result of robot error. The difference is not always readily discernable. In the case of robot error, the problem could be in the design, the hardware, or the software of the robot, meaning that the responsibility belongs to the designer, programmer, manufacturer, distributor, or retailer. Furthermore, user error may be due not just to a user's self-imposed mistake, but could be a result of poor training, erroneous instructions, or false expectations due to intentional deception.

Software responsibility is troubling since most software licenses explicitly

absolve the software developer of responsibility. A large percentage of open-source public domain software and end-user license agreements (EULA) specify that the software is provided "as is" and with no liability assumed by the developers or software companies. This includes losses of privacy or data. As privacy is a critical component of the autonomy and non-malfeasance aspects of medical ethics, such a declaration of non-responsibility is especially concerning. It is entirely possible that a software error could leak privileged information in some way, and that the software developer would feel completely within his/her rights to abdicate responsibility for such an error. From the developer's perspective, software is "take-it-or-leave-it". Additionally, a developer cannot be responsible for unforeseen consequences of every line of code, especially given that hardware updates, user error, interface and power issues, and other influences can trigger software errors. This makes the notion of responsibility extremely difficult, making the enforcement of justice related to SAR a challenging prospect, considering that software is just one of the aspects of a complete SAR system.

#### Discussion

Challenges to the core ethical principle of justice may be the most difficult to anticipate. In fact, most of the problems associated with SAR will be discovered as the robots are used in their target domains. Currently, when robots are tested in research settings with human participants, their use, distribution, and responsibility for errors, are all determined by institutional standards, and in the case of many nations, institutional review boards (IRBs). These institutions demand that the inclusion/exclusion criteria, operation of the robot, and responsibility for the robot's actions be stated in advance. Breeches to such agreements must be addressed on an individual basis, with the termination of a study as a possible consequence. However, as robots are deployed in the consumer realm, similar agreements might not be pursued.

The determination of responsibility for a socially assistive robot's actions is a

complex problem and one that must be addressed, as the technology is being developed and deployed. It is unreasonable to assume that robots will work perfectly or always be used in a completely just and honest manner. Thus, when breakdowns occur, responsibility and restitution for any harm to a user must be assessed.

### **Summary**

In this paper, we have taken the core ethical principles from medicine as a foundation for discussing ethical issues implied by socially assistive robotics technologies being developed. Since this ethical framework was constructed with ethical policies from the United States in mind, and the examples in this article come from North America and Japan, it is possible that different or additional ethical challenges arise for other cultures. More exploration is needed, especially to determine if robots designed and tested in one medical care system would behave ethically in another. Additionally, as users' reactions to robots might be different from one group to the next, proven ethical principles for one user population might not be effective for another.

New technologies bring about entirely unprecedented contexts for humanmachine interaction and call for thoughtful and well-informed multi-disciplinary studies that include inputs and expertise and address concerns from the entire complex constituency, including the technology developers, social scientists, ethicists, and, most importantly, members of the broad user community. This process must be open and ongoing, since the technologies and user responses and experiences will continue to evolve indefinitely.

#### References

- Aaronson, N. K., Acquadro, C., Alonso, J., Apolone, G., Bucquet, D., Bullinger, M., et al. (2004, Dec). International Quality of Life Assessment (IQOLA) Project.
   Quality of Life Research, 1 (5), 349-351.
- Anderson, D., Keller, J., Skubic, M., Chen, X., & He, Z. (2006). Recognizing falls from silhouettes. In Proceedings, IEEE International Conference of the Engineering in Medicine and Biology Society (pp. 6388–6391). New York, NY.
- 3. Arkin, R. (2008, March). Governing lethal behavior: embedding ethics in a hybrid deliberative/reactive robot architecture. In Proceedings of the International Conference on Human-Robot Interaction (HRI) (pp. 121–128). Amsterdam, The Netherlands.
- 4. Beauchamp, T., & Childress, J. (2001). Principles of biomedical ethics. Oxford University Press, USA.
- Feil-Seifer, D., & Matarić, M. (2005, Jul). Defining socially assistive robotics. In Proceedings of the International Conference on Rehabilitation Robotics (ICORR) (p. 465-468). Chicago, II.
- 7. Feil-Seifer, D. J., Skinner, K. M., & Matarić, M. J. (2007). Benchmarks for evaluating socially assistive robotics. Interaction Studies: Psychological Benchmarks of Human-Robot Interaction, 8 (3), 423-439.
- 9. Kahn, P. H., Ishiguro, H., Friedman, B., & Kanda, T. (2006, Sep). What is a human? Toward psychological benchmarks in the field of human-robot interaction. In IEEE Proceedings of the International Workshop on Robot and Human Interactive Communication (RO-MAN). Hatfield, UK.
- 10. Kanda, T., Ishiguro, H., Imai, M., & Ono, T. (2004, Nov). Development and evaluation of interactive humanoid robots. In Proceedings of IEEE (Special Issue

- on Human Interactive Robot for Psychological Enrichment) (Vol. 92, p. 1839-1850).
- 11. Lee, M., Forlizzi, J., Rybski, P., Crabbe, F., Chung, W., Finkle, J., et al. (2009). The Snackbot: Documenting the Design of a Robot for Long-Term Human-Robot Interaction.
- 12. Matarić, M., Eriksson, J., Feil-Seifer, D., & Winstein, C. (2007, Feb). Socially assistive robotics for post-stroke rehabilitation. Journal of NeuroEngineering and Rehabilitation, 4 (5).
- 14. Mutlu, B., & Forlizzi, J. (2008, March). Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. In Proceedings of the International Conference on Human Robot Interaction (HRI) (pp. 287–294). Amsterdam, The Netherlands: ACM New York, NY, USA.
- 16. Perelle, I., & Granville, D. (1993). Assessment of the effectiveness of a pet facilitated therapy program in a nursing home setting. Society and Animals, 1 (1), 91-100.
- 17. Robins, B., Otero, N., Ferrari, E., & Dautenhahn, K. (2007). Eliciting Requirements for a Robotic Toy for Children with Autism-Results from User Panels. In The IEEE International Symposium on Robot and Human Interactive Communication, 2007. (pp. 101–106). Jeju Island, South Korea.
- 18. Salter, T., Michaud, F., Le tourneau, D., Lee, D., & Werry, I. (2007, March). Using proprioceptive sensors for categorizing interactions. In Proc. Human-Robot Interaction (HRI). Washington, D.C.
- 19. Sharkey, N., & Sharkey, A. (2010). The crying shame of robot nannies: an ethical appraisal. Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 11.
- 20. Sung, J., Guo, L., Grinter, R., & Christensen, H. (2007). My Roomba Is Rambo:

- Intimate Home Appliances. UbiComp 2007: Ubiquitous Computing, 145–162.
- 21. Tapus, A., Tapus, C., and Matarić, M., "The Use of Socially Assistive Robots in the Design of Intelligent Cognitive Therapies for People with Dementia", Proceedings, *International Conference on Rehabilitation Robotics (ICORR-09)*, Kyoto, Japan, June 23-26, 2009.
- 22. Turkle, S. (2005, Jul). Relational artifacts/children/elders: The complexities of cybercompanions. In Toward social mechanisms of android science: A COGSCI 2005 workshop (p. 6273). Stresa, Italy.
- 23. Wada, K., Shibata, T., Saito, T., Sakamoto, K., & Tanie, K.(2003, September).

  Psychological and Social Effects of One Year Robot Assisted Activity on Elderly
  People at a Health Service Facility for the Aged. In Proceedings of the IEEE
  international conference on robotics and automation (ICRA) (pp. 2785–2790).

  Taipei, Taiwan.
- 24. Wood-Dauphinee, S. (1999, April). Assessing quality of life in clinical research: From where have we come and where are we going? Journal of Clinical Epidemiology, 52 (4), 355-363.

#### **Footnotes**

- 1 In some cases, the robot requires additional equipment to operate in an environment, such as sensors installed in the environment, or additional devices for user or operator interface with the robot. In this article, all these devices together are referred to as a SAR system.
- 2 Since autonomy can also refer to robots that are in control of their own actions, we refer to patient/user autonomy as "autonomy" while referring to the self-control of a robot as "robot autonomy" or "autonomous robots.