

# Humanoid Robots and The Social Brain: Ethical Implications

Eric Leonardis

Department of Cognitive Science  
University of California, San Diego  
9500 Gilman Drive,  
La Jolla, CA 92093-0515  
1(858) 534-6775  
ejleonar@ucsd.edu

Ayse P. Saygin

Department of Cognitive Science  
University of California, San Diego  
9500 Gilman Drive,  
La Jolla, CA 92093-0515  
1(858) 534-6775  
saygin@cogsci.ucsd.edu

## ABSTRACT

This article explores the ethical implications of introducing humanoid social robots into everyday life. Human brains are highly tuned to extract and process information from the humanoid body, and as a result, process humanoid robot bodies using the same perceptual systems. Given this specialized access to the brain, design choices regarding humanoid robot morphology could have some unintended negative consequences.

## Categories and Subject Descriptors

K.4.1 [Public Policy Issues]

## General Terms

Human Factors, Standardization, Theory, Legal Aspects.

## Keywords

Social Neuroscience, Humanoid Robotics, Ethics.

## 1. INTRODUCTION

### 1.1 Position Statement

The brain is a highly efficient learning machine that continually adapts and changes based on new information and experiences, including social interactions. Here, we discuss some potential ramifications of introducing humanoid robots into everyday human social contexts, specifically the societal and ethical implications of design decisions regarding robot morphology.

Humans are naturally oriented towards other social entities and are sensitive to social cues, even when processed and/or transmitted unconsciously [1-3]. Faces, postures, actions and body motion are critical cues for communication and cooperation; mirroring their ecological importance, these kinds of visual stimuli are processed by efficient and robust networks in the primate brain [4, 5]. Much of human cognitive and perceptual processing operates outside of conscious awareness [2, 3] but is also guided by learning [6]. Given the significance of processing humanoid features, through both phylogenetic and ontogenetic development, and that the social brain utilizes associative learning and unconscious inference, we should consider what changes might occur in the brain when humanoid robots become our social

partners [7]. Specifically, humanoid robot morphology can activate some of the same perceptual systems that are used for processing human faces, bodies and actions. Indeed, even when humans are aware that an agent is a robot, they still engage perceptual, cognitive and neural systems for social cognition [7, 8, 9]. In other words, humanoid robots, especially those that are highly humanlike in design, may have privileged access to our social neural systems, and it is possible that even if we consciously maintain a distinction between humans and robots, certain networks in our brains might not. What are some potential ethical consequences of this?

We suggest that, as humanoid robots more closely resemble humans, they can activate widespread unconscious processes such as priming, and through associative learning and neural plasticity, can lead to changes in our brains, and in turn, contribute to unintended sociocultural effects. It is known that experiences with virtual environments and avatars can produce concomitant effects in real life [10]. It is possible that social robots, through the added dimension of physical embodiment and presence, could engage human brain systems even more effectively. Below we demonstrate a select few domains in which ethical issues might arise from the integration of humanoid robots into everyday life.

## 2. QUESTIONS/ETHICAL STORY

### 2.1 Question: Body Image

There is a wealth of empirical and historical evidence that cultural representations of the body can have implicit sociocultural effects, which can impact human behavior and culture [11]. For example, young girls describe increased dissatisfaction with their bodies and may even restrict food intake after playing with Barbie dolls [12, 13]. Barbie dolls also appear to implicitly enhance gender stereotypes for career choices [14].

Humanoid robots, especially those that are close to human size and/or proportions, are more perceptually and socially salient than Barbie dolls. It is thus possible that analogous effects on body image and stereotype threat of such robots may be even greater. We suggest that potential sociocultural effects should be considered as a factor in the design of android body proportions.

### 2.2 Ethical Story: Potential Transference of Harmful Behaviors

During an ethnographic study at a humanoid robotics lab, we recorded video of a young male researcher, who during a lab tour, walked up to an android robot with female appearance, slapped its head, and said “sometimes it is therapeutic to hit the android.” Even with top-down knowledge that no physical violence against a woman has occurred, this video is generally perceived as distasteful, and as visually and emotionally confusing. Given the

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

The Emerging Policy and Ethics of Human Robot Interaction, Mar 2, 2015, Portland, OR, USA.

extremely humanlike appearance of this robot, it becomes difficult to “unsee” it as a man hitting a woman. In a situation when parts of our brains basically “see” a human, the act of the man striking the robot is perceptually a borderline case of abuse.

This example raises the issue of whether it is ethically appropriate to physically or sexually violate a robot. Even though the robot may not suffer any distress, could there be a risk of dehumanization and its transference? This can be a legitimate concern as aggression and violence appear to be reinforced rather than attenuated by cathartic practice, likely through unconscious priming and social learning [15].

### 2.3 Question: Influences on Diversity

Most androids (humanoid robots with highly humanlike appearance) in existence are modeled after young females and middle aged males with Caucasian and Asian appearance [16]. To avoid perpetuating existing issues with diversity, we need to broaden the gender, age, race and body proportions represented in the designs. In fact, robots may provide us with an opportunity to aid underrepresentation. For example, the Boston Museum of Computer Science and its “female” robotic tour guides Ada and Grace aim to strengthen the association between computer science and women, who are underrepresented as computer scientists [17]. Of course, social robotic solutions to diversity issues could be a slippery slope, and ripe sociological criticism could await those using robots as surrogates for an underrepresented minority. Empirical research is needed to assess whether robotic representation helps increase diversity or reduce bias in society. We highlight the potential of social robotics to help equity and diversity for all humans, and warn against perpetuating stereotypes (e.g., by assigning robots to a particular task due to their racial or gender morphology).

## 3. CONCLUSION

HRI as a field is socially responsible for the agents they construct and their potential influence on society. We welcome the interdisciplinary ethical discourse that has recently emerged [16]. Based on principles of human cognition and observations from anecdotal experiences, we have highlighted how robots could contribute to the emergence, maintenance, and proliferation of biases and attitudes that may negatively influence human behavior. Empirical studies based on our understanding of human social cognition and neuroscience are needed to explore psychological, social and cultural ramifications of interaction with humanoid robots, and to identify optimal design parameters.

## 4. ACKNOWLEDGEMENTS

This research was supported by the NSF (CAREER Award BCS-1151805). Thank you to Louis J. Kern for important guidance on this project.

## REFERENCES

[1] Frith, U., & Frith, C. 2010. The social brain: allowing humans to boldly go where no other species has

- been. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1537), 165-176.
- [2] Adolphs, R. 2009. The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.*, 60, 693.
- [3] Lieberman, M. D. 2007. Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.*, 58, 259-289.
- [4] Rizzolatti, G., & Craighero, L. 2004. The mirror-neuron system. *Annu. Rev. Neurosci.*, 27, 169-192.
- [5] Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annu. Rev. Neurosci.*, 31, 411.
- [6] Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. 2009. Foundations for a new science of learning. *Science*, 325(5938), 284-288.
- [7] Saygin, A.P., Chaminade, T., Ishiguro, H., Driver, J. & Frith, C. 2012. The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive Affective Neuroscience*, 7(4):413-22.
- [8] Urgen, B. A., Plank, M., Ishiguro, H., Poizner, H., & Saygin, A. P. 2013. EEG theta and Mu oscillations during perception of human and robot actions. *Frontiers in Neurorobotics*, 7.
- [9] Li, A.X., Florendo, M., Miller, L. E., Ishiguro, H., Saygin, A. P. 2015. Robot form and motion influences social attention. *Human-Robot Interaction (HRI) 2015 10<sup>th</sup> ACM/IEEE International Conference. In Press. IEEE.*
- [10] Yee, N., & Bailenson, J. 2007. The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33(3), 271-290.
- [11] Campbell, T. C., & Sitze, A. 2013. *Biopolitics: A reader*. London: Duke University Press Books.
- [12] Dittmar, H., Halliwell, E., & Ive, S. 2006. Does Barbie make girls want to be thin? The effect of experimental exposure to images of dolls on the body image of 5-to 8-year-old girls. *Developmental Psychology*, 42(2), 283.
- [13] Anschutz, D. J., & Engels, R. C. 2010. The effects of playing with thin dolls on body image and food intake in young girls. *Sex Roles*, 63(9-10), 621-630.
- [14] Sherman, A. M., & Zurbriggen, E. L. 2014. “Boys Can Be Anything”: Effect of Barbie Play on Girls’ Career Cognitions. *Sex Roles*, 70(5-6), 195-208.
- [15] Anderson, C. A., & Bushman, B. J. 2002. Human aggression. *Annu. Rev. Psych.*, 53(1), 27.
- [16] Riek, L. D., & Howard, D. 2014. A Code of Ethics for the Human-Robot Interaction Profession. *Proceedings of We Robot*.
- [17] Swartout, W., Traum, D., Artstein, R., Noren, D., Debevec, P., Bronnenkant, K., ... & White, K. 2010. Ada and Grace: Toward realistic and engaging virtual museum guides. In *Intelligent Virtual Agents*. Springer Berlin Heidelberg