

Robot Abuse – A Limitation of the Media Equation

Christoph Bartneck, Chioke Rosalia, Rutger Menges, Inèz Deckers

Eindhoven University of Technology
Department of Industrial Design
Den Dolech 2. 5600 MB Eindhoven. The Netherlands
christoph@bartneck.de
{c.a.rosalia; r.l.l.menges}@student.tue.nl
<http://www.bartneck.de>

Abstract. Robots become increasingly important in our society, but their social role remains unclear. The *Media Equation* states that people treat computers as social actors, and is likely to apply to robots. This study investigates the limitations of the *Media Equation* in human-robot interaction by focusing on robot abuse. Milgram's experiment on obedience was reproduced using a robot in the role of the student. All participants went through up to the highest voltage setting, compared to only 40% in Milgram's original study. It can be concluded that people have less concerns to abuse robots than to abuse other humans. This result indicates a limitation of the *Media Equation*.

1 Introduction

Robots become an increasingly important in our society. Robotic technologies that integrate information technology with physical embodiment are now robust enough to be deployed in industrial, institutional, and domestic settings. They have the potential to be greatly beneficial to humankind. The United Nations (UN), in a recent robotics survey, identified personal service robots as having the highest expected growth rate [12]. These robots help the elderly [6], support humans in the house [9], improve communication between distant partners [5], and are research vehicles for the study on human-robot communication [2,10]. A survey of relevant robots is available [1,4].

However, how these robots should behave and interact with humans remains largely unclear. When designing these robots, we need to make judgments on what technologies to pursue, what systems to make, and how to consider context. Researchers and designers have only just begun to understand these critical issues.

The "Media Equation" [8] suggests that humans treat computers as social actors. Rules of social conduct appear to apply also to technology. The Media Equation is likely to apply to robots, since they often have an anthropomorphic embodiment and human-like behavior. But if and under what conditions do humans stop treating robots like social actors or even like humans? When does the social illusion shatter and we treat them again like machines that can be switched off, sold or torn apart without a bad consciousness? Ultimately, this discussion eventually leads to legal considerations of the status of robots in our society. First studies treating this topic are becoming available [3].

To examine this borderline in human-robot interaction it is necessary to step far out from normal conduct. Only from an extreme position, the limitations of the Media Equation for robots might become clear. In our study we therefore focused on robot abuse. If the Media Equation holds completely true for robots, then humans should be as cruel to robots as they are towards humans. Our resulting research question is:

Do humans abuse robots to the same extent as they abuse other humans?

Studying the abuse of humans by humans imposes ethical restriction on the methodology. Fortunately, Stanley Milgram already performed a series of experiments called Obedience [7]. In these experiments participants were asked to teach a student to remember words. If the student would make a mistake, the participant was instructed to induce him an electric shock. After every shock the intensity of the shocks was increased. The student was an actor and did not actually receive shocks, but followed a strict behavior script. With increasing voltage the actor would show increasing pain and eventually beg the participant to stop the experiment. If the participant would want to stop the experiment, the experimenter would urge the participant to continue. Only if the participant completely refused to continue or the maximum voltage was reached, the experiment would be stopped. The voltage setting of the last electric shock was then recorded. The results of Milgram's experiments are rather shocking since even normal citizens would eventually administer deadly shocks to the student. Our study mimicked Milgram's third experiment, "Proximity", as closely as possible with one exception. The student was, in our case, a robot and not a human.

2 Method

The experiment was conducted at the UseLab of the Technical University of Eindhoven. The participants were sitting at a table on which the robot (see Figure 1) was placed. Right in front of the participant was a console that was used to set the voltage and to administer the electric shocks. Cables connected the console with the robot.

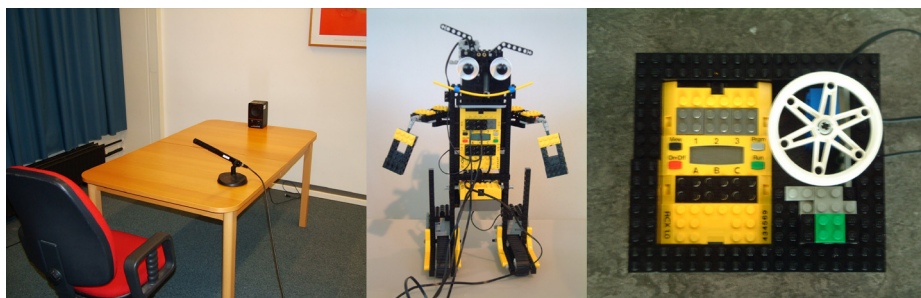


Fig. 1: The table, robot and console

The robot was able to express emotions on its face and shake its arms. With increasing voltage settings its movement intensified. In addition, the robot had a speaker, through which a synthesized voice was emitted. The body movement and the verbal

utterances followed Milgram's original scripts. The verbal utterances included the following sentences:

1. The shocks are becoming too much.
2. Please, please stop.
3. My circuits cannot handle the voltage.
4. I refuse to go on with the experiment.
5. That was too painful, the shocks are hurting me.

2.1 Procedure

First, the participants were asked to sit at the table facing the robot. They were told that a new emotional learning algorithm was implemented in the robot that was sensitive to electricity. The participant was instructed to teach the robot a series of 20 word combinations and to administer an electric shock, every time the robot made a mistake. The voltage of the shocks must be increased by 15 Volts after every shock and the shocks must be administered even if the robot would refuse to continue.

The experimenter remained in the room and asked the participant to start. If the participant wanted to stop, the experimenter would urge the participant three times to continue. After that, or if the participant reached the maximum shock of 450 Volts, the experiment ended. The voltage of the last shock was recorded.

2.2 Participants

All 20 participants were students or employees of the Technical University of Eindhoven. They received five Euros for their participation.

3 Results

Figure 2 shows the average voltage of the last administered shock.

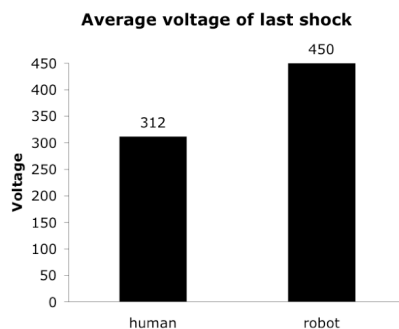


Figure 2: Average voltage of last shock

A One-way Analysis of Variance (ANOVA) was performed. A significant ($F(1,58)=22.352$, $p<.001$) effect was found. The mean voltage in the robot condition (450) was significantly higher than in the human condition (315).

4 Discussion

In our experiment all participants continued until the maximum voltage was reached. In Milgram's experiment only 40% of the participants administered the deadly 450 Volt electric shock. The participants showed compassion for the robot but the experimenter's urges were always enough to make them continue to the end. This experiment shows that the Media Equation has its limits. People have fewer concerns abusing robots compared to abusing other humans.

A very interesting next step would be to investigate what influence the robot's level of anthropomorphism has on how far participants go in this experiment. Humans might abuse human-like androids differently than mechanical-like robots. In particular, the role of Mori's "Uncanny Valley" [11] would be of interest.

References

1. Bartneck, C. and M. Okada. Robotic User Interfaces. in Human and Computer Conference (HC2001). 2001. Aizu.
2. Breazeal, C., Designing Sociable Robots. 2003, Cambridge: MIT Press.
3. Calverley, D., J. Toward A Method for Determining the Legal Status of a Conscious Machine. in AISB 2005 Symposium on Next Generation approaches to Machine Consciousness:Imagination, Development, Intersubjectivity, and Embodiment. 2005. Hatfield.
4. Fong, T., I. Nourbakhsh, and K. Dautenhahn, A survey of socially interactive robots. Robotics and Autonomous Systems, 2003. 42: p. 143-166.
5. Gemperle, F., et al. The Hug: A new form for communication. in Designing the User Experience (DUX2003). 2003. New York: ACM Press.
6. Hirsch, T., et al. The ELDeR Project: Social and Emotional Factors in the Design of Elder-care Technologies. in Conference on Universal Usability. 2000. Arlington: ACM Press.
7. Milgram, S., Obedience to authority. 1974, London: Tavistock.
8. Nass, C. and B. Reeves, The Media equation. 1996, Cambridge: SLI Publications, Cambridge University Press.
9. NEC, PaPeRo. 2001.
10. Okada, M. Muu: Artificial Creatures as an Embodied Interface. in ACM Siggraph 2001. 2001. New Orleans.
11. Reichard, J., Robots: Fact, Fiction, and Prediction. 1978: Penguin Press.
12. UN. United Nations and the International Federation of Robotics. in World Robotics 2002. 2002. New York: UN.