

Everyday robots and their potential adversarial use as harmful robots through persuasion

Kerstin Haring¹, Chad Tossell¹, Ewart de Visser¹, Victor Finomore², Kelly Satterfield³, Greg Funke³

¹United States Air Force Academy, CO, USA, kerstin.haring@usafa.edu

²Rockefeller Neuroscience Institute, West Virginia University School of Medicine

³Air Force Research Laboratory, 711th HPW/RHCPA, Wright-Patterson AFB

Abstract—This poster discusses the potential of commercially available robots as adversarial robots. Specifically, in this study we explore the potential to inflict psychological sabotage on people using persuasion. In order to demonstrate this possibility, a study using an everyday robot programmed to exert influence and designed to measure compliance is presented. We also discuss future studies of how robots are able to persuade humans to comply with certain norms, for instance, in a mission debrief context. This type of application has the potential to enhance an overall Human Machine Teaming Concept of Operations.

I. INTRODUCTION

Robots, as it is true for many machines, can be used to increase life quality and enhance human teams with their unique capabilities. It has to be kept in mind that those exact unique capabilities can pose a liability or threat to humans, originating from a wide range between carelessness to the use of robots in adversarial contexts. Misuse, the use of robots against the original intention of the designers as well as malevolent use of robots is a crucial factor that needs to be addressed through technological as well as human factors means..

II. EXPERIMENTAL STUDIES

In the experimental study, a Nao robot is programmed to prompt people to continue in a task which is undesirable to perform. The robot is trying to make people proceed beyond the time they would want to spend on that task. We are presenting the preliminary results comparing a human and a Nao robot voicing the prompts to undergraduate USAFA cadets. Results show a highly significant decrease in task time (M=35 vs. M=19 minutes) for the robot condition compared to the human condition. This means that the cadets comply substantially longer with the human than the robot. We discuss the opportunities of extending these results to robots of different size and shape. Results are expected to indicate which kind of robot design triggers more compliance and influence over people. We raise the question about the possibilities of robots of different design in order to better assist people and interact with people, as well as adjusting the design to increase their potential of persuasion.

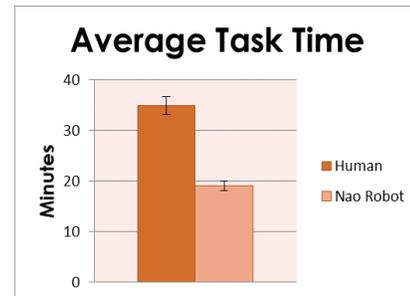


Fig. 1. The average task time of compliance with M=35 for the humans condition and M=19 minutes for the Nao robot condition.



Fig. 2. The current conditions of this study with a human control and the Nao robot, and the future conditions using a Baxter robots and a modified Roomba for experimental comparison.

Future studies need to take into account the social application of robots in mission planning scenarios. To address this current lack in research, we use virtual agents and robots to resolve a conflict introduced in a mission. Our objective is to identify robust strategies of machine involvement in conflict detection and resolution within mixed human-machine teams. This will extend current research showing the potential of agents to be persuasive, convey information, and potentially resolve conflict [1] with the purpose of application in a military or search and rescue mission.

ACKNOWLEDGMENT

We would like to thank 2nd Lt Sarah Pruznick for her extensive work in data collection, as well as 2nd Lts Julie Fleming and Chris Unger. This work is supported by the Air Force Office of Scientific Research under Award Number 16RT0881 under Program Officer Benjamin Knott.

REFERENCES

- [1] Shen, S., Slovak, P., & Jung, M. F. (2018, February). Stop. I See a Conflict Happening.: A Robot Mediator for Young Children's Interpersonal Conflict Resolution. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (pp. 69-77). ACM.
- [2] Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism
- [3] increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113-117.
- [4] Milgram, S. (1963). Behavioral Study of obedience. *The Journal of abnormal and social psychology*, 67(4), 371.
- [5] M. Siegel, M., C. Breazeal, and M.I. Norton. "Persuasive Robotics: The influence of robot gender on human behavior." *Intelligent Robots and Systems*, 2009. IROS 2009. IEEE/RSJ International Conference on. 2009, 2563-2568.