COLLEGE OF ENGINEERING

THE THEATER METHOD: EXPLORING UNETHICAL RESEARCH TOPICS IN HRI SAMARENDRA HEDAOO, DR. HEATHER KNIGHT, OREGON STATE UNIVERSITY

Inspired by acting, the Theater Method allows research participants to have in-person experiences of a robot doing unethical actions, for example, violating their character's privacy[1] without being emotionally damaged themselves.

GENERATIVE SCRIPTS ENABLE EXPLORATION OF MANY VARIABLES

- The Theater Method explores many research variables by varying a script between two humans and a robot, evaluating variations via a large-scale video study, and a smaller-scale in person study.
- It has benefits like traditional user studies for new research topic areas but provides greater psychological and informational safety to its participants[2][3] because the violations are simulated.
- We believe that it will be particularly helpful when designing new social functionality into machines, and in areas where user sensitivities are not yet known.





Electrical Engineering and Computer Science



The robot was polite. Agree Somewhat agree Neither Somewhat disagree Disagree

recombinant films for online study (audience perspective)

ACTOR PERSPECTIVE -> EMOTIONAL REACTIONS & EXPLANATION The actor perspective, where the participants had scripted interaction with a robot and a confederate (below), provoked quite a lot of emotions from our participants.

printed scripts for in-person study (actor perspective)

FUTURE WORK: SEEKING UNETHICAL RESEARCH TOPICS

Our initial exploration of the Theater Method considered a robot barista commenting on its clients, revealing what conversational robot data use participants found appropriate. In future work, we would like to consider other research topics of high psychological risk, offering insight to robot programming and how future regulations of social robots should be structured.

AUDIENCE PERSPECTIVE -> LARGE SCALE VARIABLE EXPLORATION

The audience perspective enables large-scale data-collection; in its first use, we used Amazon Mechanical Turk. Participants viewed videos of a robot commenting on its customers conversations and answered single question surveys, such as:

The robot was: polite/impolite appropriate/inappropriate *considerate/inconsiderate* privacy-protecting/privacy-violating





Experimenter Participant

[1] "A Robot Barista Comments on its Clients: Social Attitudes Toward Robot Data Use" in 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI) [2] R. C. Andres-Hyman, J. S. Strauss, and L. Davidson, "Beyond parallel play: Science befriending the art of method acting to advance healing relationships." Psychotherapy: Theory, Research, Practice, Training, vol. 44, no. 1, p. 78, 2007. [3]A. Blatner, "Acting-in: Practical applications of psychodramatic methods." Springer Publishing Company, 1996

