

The Theater Method: Exploring Unethical Research Topics in Human-Robot Interaction

Samarendra Hedao, Dr. Heather Knight
Oregon State University, Corvallis, OR, USA
{hedaoos, knighth}@oregonstate.edu

The Theater Method explores many research variables at a time using the same script structure repeated many times with small variations based on the research variables. 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 without actually being emotionally damaged themselves. As an accoutrement to the in-person actor perspective, we can also collect data about the audience perspective via online video study, which also helps identify research variables for in-person exploration.

The Theater Method has benefits similar to traditional user studies for new research topic areas, but provides greater psychological and informational safety to its participants [1] [2] because the violations are simulated. Previous methods exploring sensitive topics like people's privacy expectations of a robot are often at a distance from the privacy-violation (survey, video studies [3]); conservative (user study), or at a danger of putting the participant at risk (live deployment).

In our first experiment utilizing the Theater Method, we explored people's attitudes toward robot data use in two phases: (I) where the participants see a robot interaction from an audience perspective and then (II) where the participants get to experience interaction with the robot from an actor's perspective as they act in the same scene as (I) along with the robot and the same professional actor from (I). At the end of each scene in (II), the participants complete a survey.

In varying the many scripts, we were able to explore themes of the robot's data use, how it used that data, whether the comment it made was positive or negative, to whom it addressed the information. The audience perspective is great for collecting large-scale data about what might matter in an interaction, while the actor perspective provoked quite a lot of emotions from our participants to a range of comments by the robot.

In future work, we would like to apply the Theater Method to other sensitive questions in robotics, such as a robot's moral decision-making (the trolley problem is not something that should be evaluated live), or a robot's potential role in mediating workplace harassment and civil interactions. We believe that the Theater Method will be particularly helpful when designing new social functionality into machines, and in areas where user sensitivities are not yet known.

Citations:

[1] 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.

[2] A. Blatner, "Acting-in: Practical applications of psychodramatic methods." Springer Publishing Company, 1996

[3] S. N. Woods, M. L. Walters, K. L. Koay, and K. Dautenhahn, "Methodological issues in HRI: A comparison of live and video-based methods in robot to human approach direction trials," in ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2006, pp. 51-58.

[4] Hedao, S., Williams, A., Fallatah, A., Wadgaonkar, C., & Knight, H. (2019) "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)