PURDUE COLLEGE OF SCIENCE – COMPUTER SCIENCE DEPARTMENT

EVALUATING ROBOTS' SOCIAL INTELLIGENCE

Presenters: Avery Kruppe and Denison Guvenoz || Mentors: Dr. Sooyeon Jeong and Dong Won Lee

PURDUE UNIVERSITY

mit media lab

OVERVIEW

Robots, coupled with AI, have been taking on more of a human likeness in their interactions. However, robotic assistants like Alexa or Google Home are not yet capable of holding meaningful, socially normative conversations. This project aims to evaluate and improve "social intelligence" in robots so they can act as more effective companions.

DEFINING SOCIAL INTELLIGENCE

Effectiveness is defined as a robot's social competencies compared to errors. Competencies result when the robot acts according to social conventions, and errors occur when the robot's actions deviate from these (Tian & Oviatt, 2021). We evaluate social intelligence both in terms of the robot's performance and socio-affective instances.

"Performance" instances relate to the robot's mechanical and technical functions. It includes how fast the robot responds and if it understands the user's words.

"Socio-affective" instances relate to how the robot recognizes and responds to the participant's overall social behavior. Categories include:

- **Conversational Mechanics:** Conversational flow, specifically when the robot responds at the right time, often after the user.
- **Social Norms:** The robot's ability to act according to what is considered polite and appropriate; for example, thanking the user for their time at the end of a session.
- **Engagement:** Recognizing and addressing whether or not the robot has the user's attention; Attempts should be made to assess the user's interest in activities and grasp the attention of the user.
- **Emotions:** Recognizing and responding to the user's feelings, expressed verbally or nonverbally. The robot should be able to sympathize and reflect appropriate emotions back.
- **Social Contexts and Relationships:** The robot understanding the role it plays in contrast to the user and using correct language that reflects that. It also relates to acting in line with the social setting.
- **Understanding the User's Intent:** Foresight when recognizing a user's intentions, supporting the user's intents, and adapting the conversation around them; for example, when the robot understands that the user is ready to continue the conversation.
- **Understanding Knowledge State:** The robot's ability to recall previously learned information and apply it; for example, addressing the user by name.

METHODOLOGY

To evaluate effectiveness, we are annotating videos of 70+ clinical participants' therapy sessions facilitated by Jibo, a social robot turned mental health assistant. We note Jibo's errors and competencies, watching how these affect the user short or long term, and consider how Jibo's interactions can be improved by describing recovery behavior as needed.

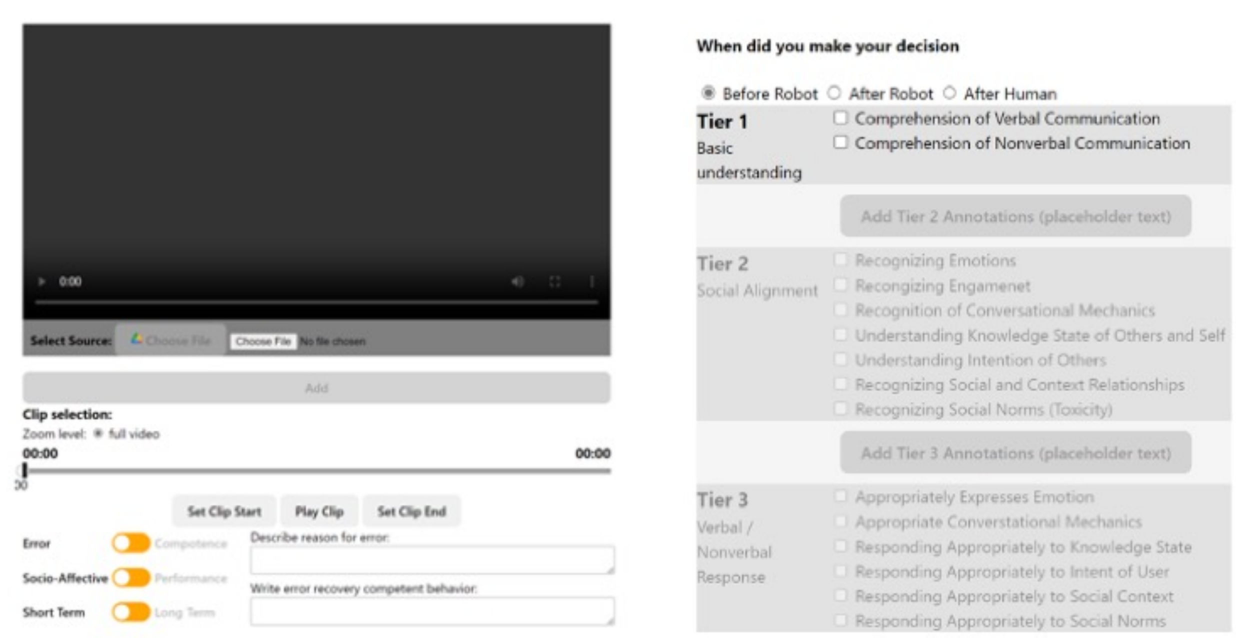

Figure 1: *The annotation tool used to analyze the videos of Jibo's interactions with participants. Videos are viewed on the left, and the social attribute selections are made on the right.*

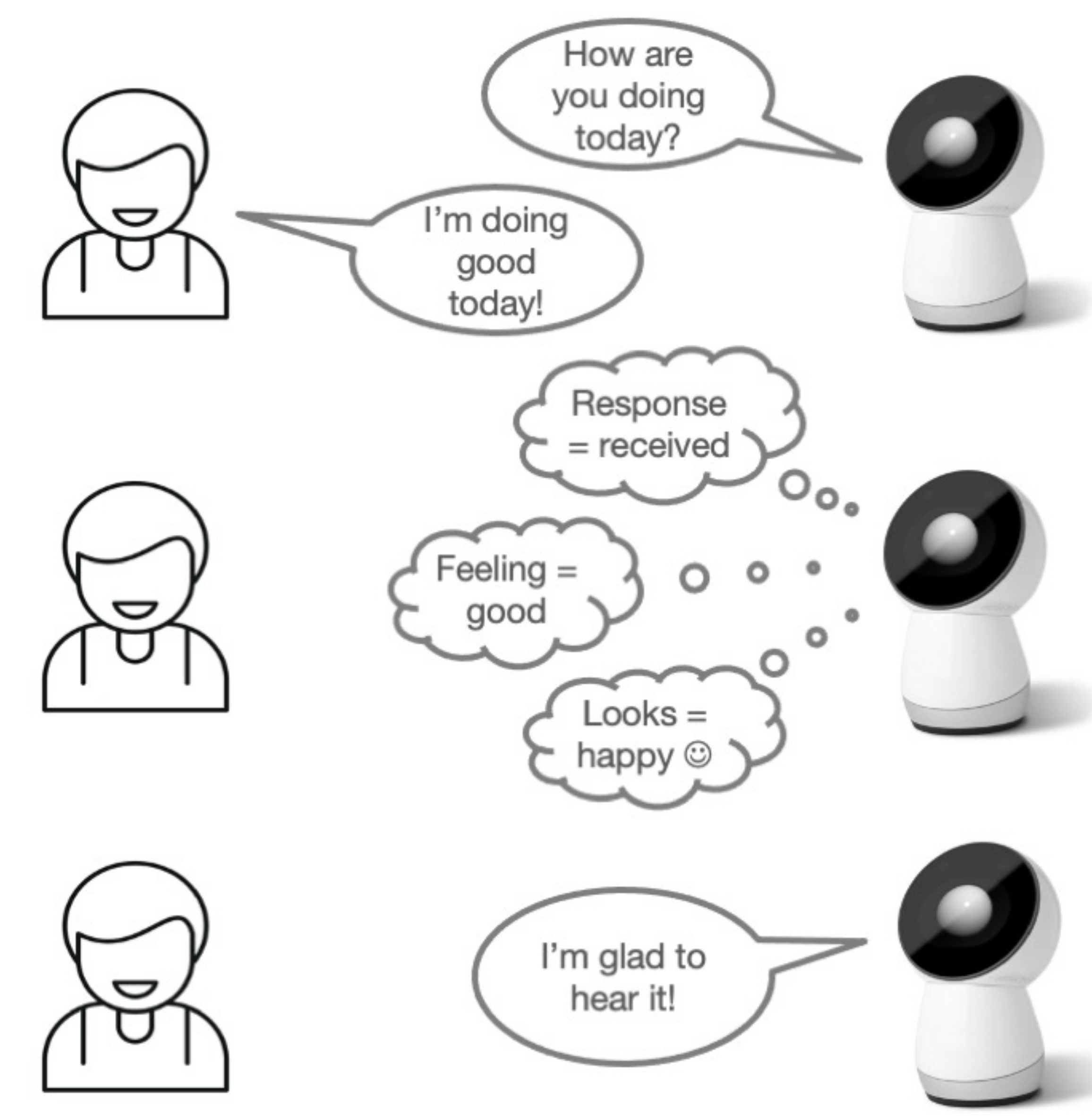

Figure 2: *Jibo communicates with participants by observing how they respond and act. In the example above, Jibo listens for verbal cues such as emotions and interprets the participant's nonverbal reactions such as smiling. Using these, it develops an appropriate response.*

DISTRIBUTIONS OF ATTRIBUTES ACROSS VIDEOS

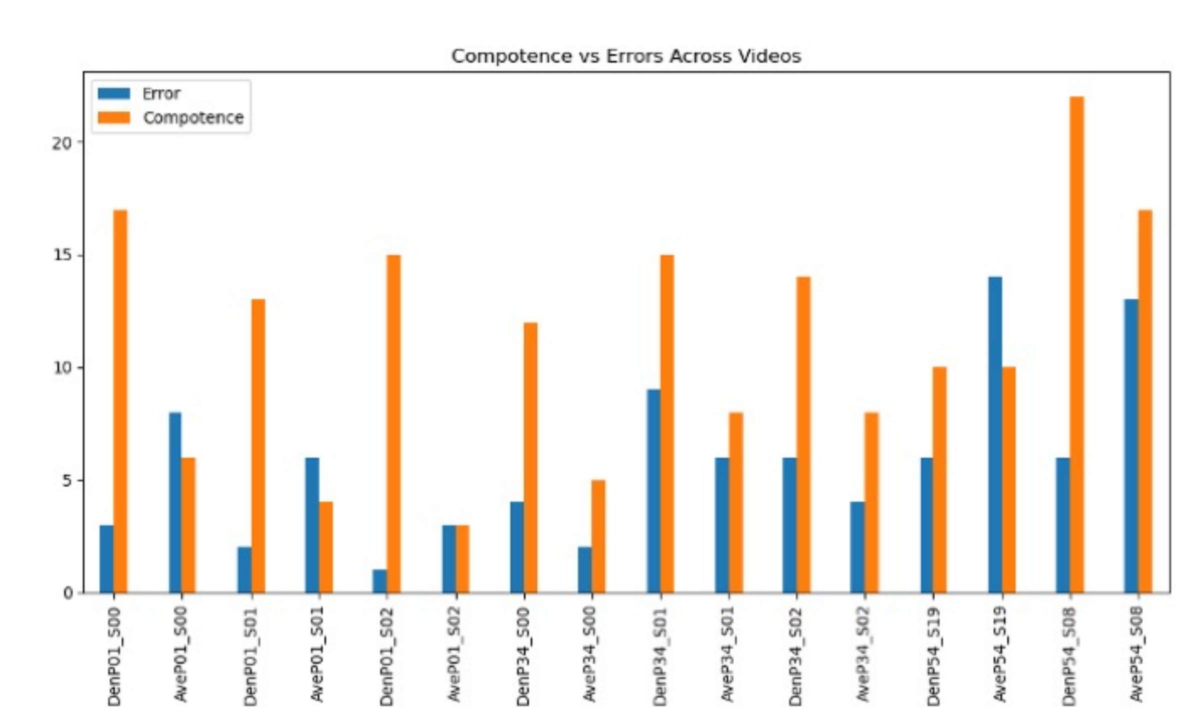

Figure 3: The number of "Social Competence" and "Error" instances are noted. The counts are tallied for each video annotated. Disagreements on what is deemed competence versus error were abundant. Sometimes an annotator would note one instance whereas the other would not notice that specific interaction, leading to varying quantities of annotations.

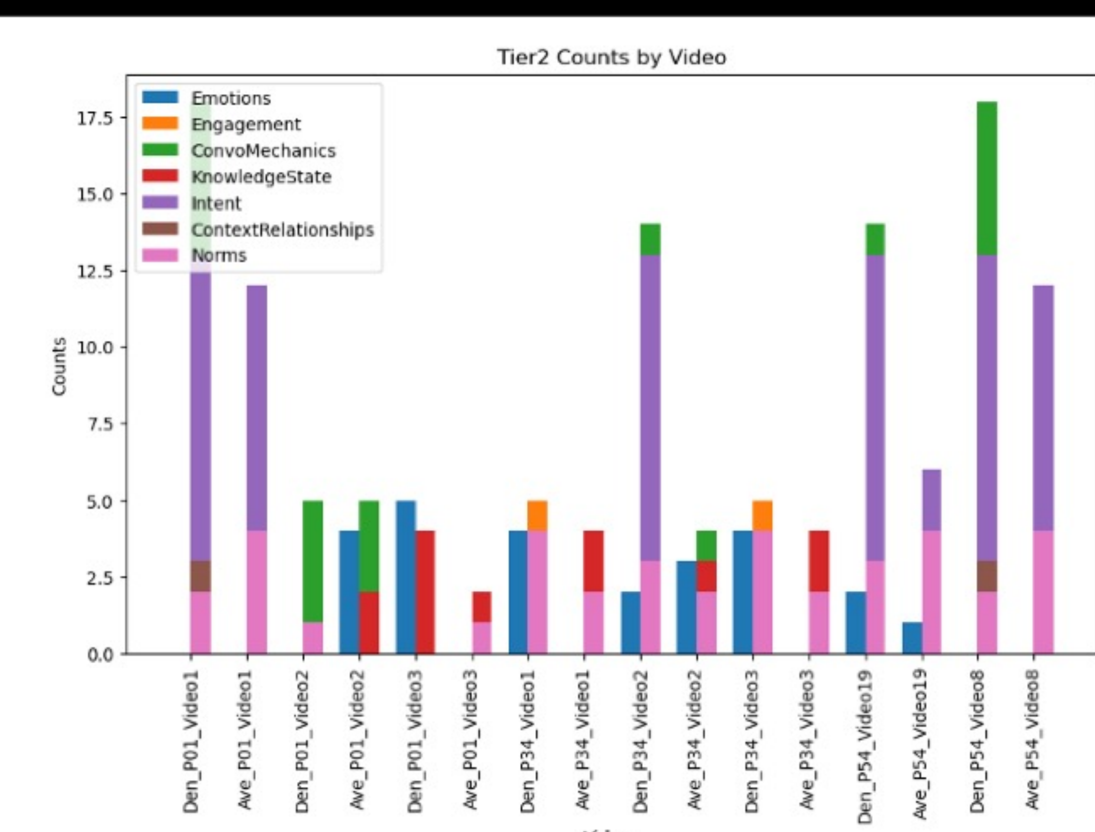

Figure 4: The number of specific attributes of communication (emotions, engagement, knowledge of conversational mechanics, etc.) are noted. The counts are tallied for each video annotated. This figure is more detailed than Figure 1, with a look at how annotators interpreted the various social intelligence attributes differently, even if the same exact interaction was noted.

MAIN FINDINGS

As shown in the figures, "social intelligence" varies from person to person. Categories such as social norms and social contexts, for example, are heavily subjective and based on learned experiences.

Thus, our current conclusions go beyond the participant and robot interactions and center around the bias involved in social contexts. **Socio-affective instances can be categorized, but each category is largely subjective.** Differing opinions in our annotations support this, and they pose an interesting question on whether robots can be taught the social norms needed to perform effectively. This question, along with further annotations and analyses, will help us understand the potential that robots, including Jibo, have as sympathetic companions.

NEXT STEPS

It is our goal to see if we can standardize the definitions for the socio-affective categories, get more similar annotations between ourselves, and educate future annotators on what to look for. Our data will increase as more researchers join us in annotating socio-affective instances. Once we have sufficient data, we hope to use this to teach our robot how to best embody the categories. **We hope that a diverse collection of annotators' ideas can help make our robot socially intelligent by adapting to a wide range of social expectations.** We hope our data on Jibo's effectiveness can teach us how to improve other social robots' intelligence as well, moving forward.