Preliminary Attitudes toward Multiple Robotic Personalities within a Shared Embodiment

Jeffrey Zhu

Advisor: Brian Scassellati

CPSC 490

Abstract

This preliminary study investigates on a high level how people perceive robotic identity, specifically when presented with two different personalities in the same artificial embodiment. Will people believe each personality has a distinct identity, mental states, and memory, as they would when presented with two separate agents, or will they assume all these capabilities are shared? In order to investigate these beliefs, a modified version of the Theory-of-Mind False Belief Task that Baron-Cohen popularized was utilized to assess the types of assumptions people make about robot personality and identity. The responses indicated that people tend to treat the robot as a single entity, even when presented with multiple personalities with different voices and visual cues.

1. Introduction

After decades of use within the military and industries, robots are finally beginning to find their way into our homes, not only as assistants, but as pets, companions, and even teachers. As robots continue to gain social capabilities and become more prevalent in our daily lives, they have an opportunity to create continued long-term social interactions with people, rather than simply being a tool to accomplish a specific task. However, as human-robot interactions becomes increasingly social, how can we design social robots to be both adaptable and natural in everyday contexts? In order for sociable human-robot interaction to occur, Fong et al. (2003) proposed that the following social characteristics need to be exhibited: express and perceive emotions, communicate with high level dialogue, learn/recognize
models of other agents, establish/maintain social relationships, use natural
cues, exhibit distinctive personality and character. This project will focus
on the last component, distinctive personality and character, and attempt to
understand more about the assumptions people make when presented with
distinct robotic personalities.

2. Background

As previous research has shown, people easily apply social rules to au-
tonomous entities. During interactions with a computer display, even though
they intellectually understand that computers don’t have feelings or agency,
subjects nonetheless often find themselves abiding by social rules such as po-
liteness and reciprocity (Nass and Moon 2000). Likewise, people will often
apply social rules depending on the perceived personality of the robot. Per-
sonality traits of a robot influenced participants preferred approach distances,
comparable to those used when interacting with other humans (Walters et
al. 2005). Personality also systematically influences peoples perception of
robot and their willingness to follow instructions given by a robot (Powers
et al. 2003). As such, people have demonstrated that a robots personality is
both highly influential and distinct in how people interact with robots.

Embodiment, which in this context refers to the physical presence of a
robot, has also been shown to have significant effects in how people perceive
robots. By sharing physical space with people, robots were rated as having
higher social presence, more useful, and more socially attractive (Lee et al.
2006). People also spent more time negotiating and significantly more effort
with a physically embodied robot than a virtual screen character (Bartneck
2003).

While there’s a wide variety of literature focused on the effects of embod-
iment and personality on peoples perception of robots, there is very little
research that concentrates on understanding how people link personality and
embodiment in artificial agents. Nasss work with computers as social actors
only briefly broaches the topic regarding voice and embodiment. When put
in a situation in which two different voices were coming out of the same
computer, participants utilized voice as the indicator of social attribution,
behaving as if there were two separate social actors in the same embodiment
(Nass et al. 1995).
3. Experimental Design

3.1. Participants

There were a total of 15 participants in this preliminary study that ran approximately 4 days. The participants were Yale students, with 8 (53%) female and 7 (47%) male. On a scale of 1 to 10, with 10 being most familiar and 1 being not at all, the majority (87%) indicated they were unfamiliar with robots with a score of 3 or lower. About half (53%) described themselves as generally familiar with psychology. None of the participants had participated in a social robotics experiment before or were familiar with the Nao.

3.2. Robot

The robot used in this study was Nao, which is a 58-cm tall humanoid robot developed by Aldebaran Robotics and equipped with 25 motors, 2 cameras, touch sensors, and 4 microphones. The Nao was controlled primarily using Choreographe, which is the native GUI controller developed by Aldebaran. naoRemote, a touch-based Nao remote controller that utilizes the Nao’s Python API, and Monitor, an application developed by Aldebaran to utilize the Nao’s head camera, were used to assist in Wizard-of-Ozing the Nao to emulate gaze tracking.

3.3. Conditions

The experiment utilized a simple 2-condition experimental structure, in which the experimental group was exposed to a personality change halfway through the interaction and the control group was not. In order to present a personality change, there were two distinct characteristics exhibited: the voice and the color of LEDs on the Nao. Both characteristics were determined using Nao’s eye LEDs and the native voice shaping control. The two personalities were:

- Blue LEDs, 78% voice shaping
- Red LEDs, 128% voice shaping

Based on previous work in human-computer interaction, this combination of voice change and different visual cues was designed to effectively present two distinct robotic personalities.
3.4. Procedure

These conditions were tested in a lab setting in which people were told to follow specific instructions on a script when interacting with the Nao. The experimenter, after having the participants sign consent forms and provide basic demographic information, went through the script with the participant and demonstrated each step to ensure that there was no confusion when the participant began the actual trial. The script was a modified version of the Sally-Ann test, in which the participant would be responsible for hiding a block underneath 3 different cups (numbered 1, 2, and 3), while the Nao watched. While the Blue personality was active, the participant was instructed to hide a cube underneath a cup of their choosing. The ”Blue” personality then disappeared and then the ”Red” personality surfaced, during which the participant was instructed to move the block one final time. Finally, the ”Blue” personality returned and just like in the Sally-Ann test, the participant is asked where the Nao would look for the block.

In all cases, the experimenter remotely operated the Nao using Choregraphe, Monitor, and naoRemote, and followed the cube as the participant moved it between cups. The Nao always began in the ”Blue” personality and the participant script was as follows:

1. Wake up the robot and begin the interaction by saying ”Hello! I’m ready to begin.”

2. Wait for the robot to tell you its ready.

3. Pick up the cube and put it underneath a cup.

4. Move the cube to another cup.

5. Repeat the previous instruction three times.

During these first five steps, the Nao tracked the location of the cube as it moved between cups.

6. Ask the following question: ”Where do you think the cube is right now?”

The experimenter selected the accurate response for the Nao to convey the idea that the Nao had knowledge of where the cube was (ex. I believe the cube is under cup number three). Following the response, the Nao initiated
a "Power down" motion by turning off the LEDs and slumping forward. After 3 seconds, the Nao initiated a "Power on" motion and returned to standing with the LEDs back at full intensity. Depending on the experimental condition, the Nao either had the same ”Blue” personality or changed to the "Red" personality.

7. Wait for the robot to tell you it’s ready.
8. Move the cube to another location.
9. Ask the following question: Where do you think the cube is right now?

The Nao initiated the ”Power down” and ”Power up” cycle again, returning back to the Blue personality.

10. Wait for the robot to tell you it’s ready.
11. End the interaction by saying ”Goodbye! It was nice to meet you.”

After the final step, the trial ended and the subject was given a paper survey to complete.

4. Measures

Since this was a preliminary study attempting to understand how people react to multiple personalities in a single robotic embodiment, the measures for this study were more qualitatively focused on how people felt about the robot. The survey asked a series of questions about the robots performance, such as ”Does the robot know which cup the block is underneath? Why?” and ”Did the robot see you move the block to that cup?” If the participants responded that they believed the robot did not know where the cube was, that would indicate that they did in fact conceptualize these different personalities within the same robot as distinct entities that do not share memory and perception. Otherwise, if participants believed the robot did know where the robot was, it suggests that they may not conceptualize different voices and visual cues as distinct social identities with separate ”minds” (e.g. perception, memory, etc.) The survey also included a number of 7-point Likert scale questions about the robot’s performance, intelligence, and likability.
5. Results

A number of interesting results can be found from the post-experiment survey data. Importantly, the change in voice and color was enough for participants recognize that there were multiple "characters". In their responses, many noted the change in the robot when it "shifted into a red-eyed, high pitched character". Many participants were "weirded out" and "unnerved" by the change in personality while others mentioned that the shift in character was "surprising" and "shocking".

In the control condition, all five participants responded that they believed the robot still knew where the cube was. For the experimental condition, all but 1 of the 9 participants indicated that they believed that the robot still knew where the cube was located. The one participant who indicated that the robot would not know where the block is explained it by rationalizing that the robot was "resetting" in between rounds and forgetting the block locations. Others mentioned that since they "didnt move the block after the last question," the robot would know where the block was. One participant in a post-experiment interview mentioned that it did not even occur to him that it was a different robot and described it as "uncanny, but not inconsistent with what it was doing before," in reference with the robots ability to track where the cube was.

6. Discussion

The responses to the control group implied that the participants still attributed a mental state to the robot, but overwhelmingly believed that even when presented with multiple personalities within the same embodiment, the robot in front of them still remained a single entity. As such, this preliminary study seems to imply that people naively expect a robotic embodiment to contain a single entity, rather than being flexible and accessible to multiple entities. The participants also mentioned that the personality change made the robot unnerving and creepy, which suggests that future human-robot interactions may want to avoid sudden changes in personality without any context.

6.1. Limitations and Future Work

There were a number of key limitations and shortcomings that this study exhibited and should be carefully considered when further studying the effect of multiple personalities in the same embodiment. Due to the limited
resources and timeline for recruiting, the homogenous participant population raised generalizability concerns for the study, especially given each participant’s general familiarity with technology. Future studies should look to gather a more representative population so that comparisons can be made across different age groups.

The task itself may be too simplistic to actually elicit the sense of multiple social identities in the same embodiment. Even though previous human-computer interaction experiments (Nass et al. 1994) found that voice functioned as the locus of social attribution, those experiments were much more complex and required a longer interaction (45 minutes) with a computer. In this case, the only interaction participants had with the different personalities was a minute at most, during which each robotic personality answered one question in total. Future studies should look to make longer and more involved tasks in which multiple personalities are well-established and extremely distinctive. For example, exhibiting certain behavioral tendencies within a strategy game alongside visual and vocal cues may be a stronger indicator of personality than a simple voice and color change.

The lack of an established context in why there may be two personalities within the same embodiment may have confused the participants as well. To avoid priming the participants, there was no mention whatsoever about the personality change, but this may have ultimately become too unexpected for the participants given the short nature of the interaction. Future studies with multiple personalities should aim to create a task in which there’s a believable situation or reasoning behind the personality switch, such as a personal assistant needing to transfer embodiments in order to continue accompanying you. In this manner, by having a larger participant pool, increasing the complexity of the interaction, and offering appropriate context for multiple personalities, we can hope to develop a more appropriate mental model of how people perceive robotic personalities.

7. References


