Social Robots: Mimicking Human Gestures

Peter E. Jacobo

May 10, 2017

Abstract

Social robots are designed to interact with people. Due to physical limitations they sometimes require human assistance to meet their goals. In this experiment SARAH, a social robot, invites people in her environment to ask her a set of questions. For SARAH to succeed she must be able to engage someone in a short conversation and have them assist with a task. A dynamic social robot could take advantage of nonverbal cues such as gestures may be more successful in interacting with a person than a static robot that does not take advantage of nonverbal cues. SARAH has been programmed with a set of gestures made to mimic common gestures used by people in conversations. This experiment investigates whether the use of the gestures influences people's perception of SARAH's intelligence, naturalness. Participants are also invited to describe SARAH in their own words.

Contents

1 Introduction

Social robots are robots that interact with humans and exhibit anthropomorphic qualities. They might not always look like humans but they strive to follow social rules as if they were a person. Adhering to social rules is what allows them to work alongside people. However, due to current physical limitations robots often need help. A robot with no arms and a simple moving mechanism will not be able to accomplish some tasks such as changing floors. Some social robots can climb stairs [4] but for most, alternative strategies must be found. Seeking help is an essential social robot ability.

Due to the reliance on human-robot interaction, social robots need to be able to navigate through complex social norms that are difficult to define and therefore encode into software. One of the first considerations for a social robot is its body. Research has shown that people are more comfortable around social robot bodies with anthropomorphic qualities. The participants comfort level is important to consider because it impacts their willingness to complete a task that involves being around a social robot. Also, if people can visually see that the social robot is intended to look humanoid then they can more easily understand that its behavior is also attempting to seem human-like. While appearance and behavior are separate in implementation, they blend together into the experience that someone has when interacting with a social robot.

Traditional robots complete tasks that can be easily represented graphically. While those tasks may be complex, there is a general understanding of the environment. In contrast, social robots are used in environments where there can be many interruptions. Human behavior can be difficult to predict.Therefore social robots must be able to adjust and respond to changing situations[2]. This flexibility comes naturally to humans and is much more difficult to implement in a mechanical being. However there are techniques that can make a social robot more adept in its setting.

One strategy to encode responsiveness in social robots is to augment their communication abilities beyond dialogue. This may require the social robot's physical form to meet certain hardware conditions - a monitor or physical face is needed to show facial expressions and a moving base, or even legs, are necessary for movement. A social robot's hardware limits its ability to express beyond dialogue in some situations, but there are other features - non-word utterances for example - that can be used without adding hardware. The goal of this project is to add facial expressions, gestures, and non-word utterances to a social robot in order to observe their effect on a conversation between the social robot and a person.

The social robot used in this project, nicknamed SARAH, has no arms and a moving base - a pioneer P3DX robot. SARAH also has a monitor, laser scanner, kinect sensor, and keypad. While the absence of hands and legs limit gesture capability, there is much room for other types of non-dialogue communication, even gestures. In this project SARAH has been programmed to perform facial expressions, gestures - these are based on the movement of the P3DX base - and also non-word utterances, that is to say sounds such as "mhm" and "hmm.." that represent a form of verbal expression without words. These three elements, referred to as features, are activated by an experimenter using a graphical user interface (GUI). In the GUI, the features may be implemented in isolation or in combination. The main idea of the experiment in this project is for SARAH to have a conversation with a participant while a researcher inserts features into the dialogue.

The experiment takes place "in the wild" which is to say a non-laboratory setting. This choice stems from the idea that a human-robot interaction observed in a natural setting would have more internal validity than one in an artificial environment. This also increases the confidence that results or observations made from the experiments may be generalized to other human-robot interactions in non-lab settings. That experiment data will be gathered in two forms. The conversations will be recorded on video and there will be a survey for participants to complete after the conversation with SARAH is complete. From both sources of information, the goal is to observe how the features, in different combinations, affect the clarity of SARAH's communication with the person and also the person's comfort level speaking and working with a social robot.

2 Background and Related Work

2.1 Social Perception of Robots

For many people, interacting with a social robot is a novel experience. Robots have been around for decades making daily life easier, however social robots with anthropomorphic qualities are comparatively new. Fussell et al. believed that people tend to attribute "social categories" to social robots based on their interactions with them[3]. People make assumptions about other people based on first impressions and visible characteristics and this also applies to people encountering a social robot for the first time. In order for a social robot to succeed in an environment with people negative perceptions must be minimized. Negative outcomes of a person meeting a social robot include fearing the robot, hesitating to interact with the robot, and not understanding that the social robot can be interacted with as if it were a person. According to the study conducted by Fussell et al. people perceive a robot's behavior separately from its personality. While many attribute anthropomorphic qualities to the robot's behavior people believed that the personality was "mechanical" and therefore not capable of feeling emotions.

According to research by Carpinella et al. there is another dichotomy that can be used to describe people's perception of other people[1]. Their findings are that people can be described by a warmth component and a competence component. The warmth component serves as a measurement if that person's demeanor and it can range from friendly to hostile. The competence component describes the person's ability to carry out actions from being productive, if they have a high warmth component, to be a threat. Carpinella et al. believe that this abstract representation of people's perception of others can also be applied to people's perception of a social robot. The goal for a social robot is for them to be perceived as high on the warmth scale, however a bit low on the competence scale. The reason for this is that social robots often demand help from people. According to Carpinella at all those with a high warmth and low competence evoke feelings of pity from the person. These feelings can push a person to help a social robot meet its needs.

2.2 Facial Expressions

Many social robots are able to communicate with people through dialogue however there are many benefits for using nonverbal cues such as facial expressions. Facial expressions that occur on their own or in combination with speech and other nonverbal cues. Menne and Lugrin believe that facial expressions communicate emotions during conversations[7]. With facial expressions robots are able to express their emotions more naturally. It would be very unnatural for a robot to always state their emotions through dialogue (i.e. "I feel happy", "I am sad") but if the social robot instead used a facial expression the person could infer the emotion and relate it to the conversation through context. Facial expressions help clarify the meaning of what a social robot says or does and also serve to simulate empathy in the social robot.

Designing facial expressions requires a lot of thought towards the different components. Most social robots with faces have at the very least eyebrows, eyes, and a mouth. In an experiment where participants looked at different cartoon facial expressions Koda et al. made observations about a link between certain parts of a face and how easily those parts were able to convey emotion as determined by the participants who attempted to identify the correct emotions[6]. For their participants the mouth was the component that was most telling of the emotion being expressed. The eyes were also significant but the level of significance varied depending on whether the participant's background was Japanese or Hungarian.

2.3 Personality

Social robots can be thought of as having personalities. These personalities are primarily based around behavior and physical features. Jung et al. found that people who are extroverted are more fond of extroverted social robots and those who are introverted are more fond of introverted social robots[5]. It seems logical that people are most comfortable around those that are like themselves. This provides a challenge when designing the personality for a social robot. If the personality is designed to be strictly introverted or extroverted than people of the opposite personality may not feel as much friendliness towards the robot as they may have been otherwise. A social robot that can switch between the two personality types may be challenging to design, and it would also require the experimenter to correctly identify which category

the person falls into. It may not be enough to observe their interaction with the robot. However, if a robot is meant to be in a public area attracting people's attention, then it seems most appropriate for the robot to lean on the side of being extroverted. The social robot must be extroverted enough to engage people in conversation inviting them to help it while not making them feel uncomfortable.

3 Methods and Design

3.1 SARAH the Social Robot

In this experiment I use SARAH the social robot. Her name acronym stands for Socially Acceptable Robot that Asks for Help. SARAH has a moving base that is itself a Pioneer P3-DX robot. With its wheels it is able to turn in place and move at the speed of a person walking. SARAH also has a tablet that acts as her display. It can be used to display a face such as the robot face program and serves as the main point of interaction with people. On top of the tablet is a web cam that is used to record interactions with participants. SARAH has on-board speakers for outputting sound. This also her to talk with people and with the web cam we can also hear replies in real time. Like several social robots SARAH is limited by not having hands or the ability to climb stairs. However, this leads to using SARAH in ways that have her interact with people.

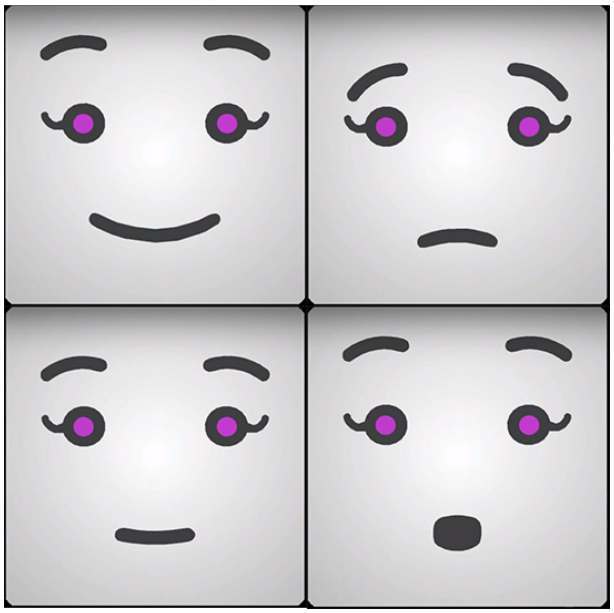

Figure 1: The four facial expressions that SARAH uses in the experiment

Shown above are the facial expression that are used in the experiment. The facial expressions are generated by the robot face program. The expressions are identified as **smiling**(top-left), **frowning**(top-right), **neutral**(bottom-left), and **surprised**(bottom-right). The main factor contributing to each of SARAH's facial expressions is the shape and orientation of the mouth. The eyebrows also change to accentuate the mouth. SARAH's eyes do not change throughout the different facial expressions however they do have eyelashes, suggesting her female gender. This robot face is designed to look friendly and approachable. The facial expressions chosen for the experiment provide a wide range of emotional expression without including emotions that could cause stress to the participants, such as a anger.

3.2 Location

The experiment task takes place in the hallway between the Olin and Wold academic buildings. This location is key because it is an area with a large amount of traffic but since there are fewer lounging areas

than other areas, there is also less noise. The volume of the environment is very important when selecting the area because SARAH can not adjust the volume of her voice after being placed in the environment. Another consideration that lead to choosing this location is visibility. SARAH is easily visible by people passing through the hallway from either direction, she can also be seen by those descending the stairs next to her. The main drawback to the location has been the amount of directions from which people can walk towards SARAH. While SARAH is easily visible, she can not view her entire environment at the same time. During experimentation SARAH faces the entrance to Wold and from the position she also has a clear view of the stairs towards Wold auditorium and people coming around the corner from the direction of the Crochet lab. However this leaves her back facing Olin and it is more difficult for the researcher to identify that there is someone approaching from that direction. This problem is partially due to SARAH's location and partially due to the technology used for viewing the environment. SARAH's front facing web cam can only see what is in front of her. To resolve this issue, in future iterations of this experiment environmental cameras would greatly compliment the camera on-board SARAH.

Figure 2: A student interacts with SARAH in the experiment location

Figure 2 shows SARAH in profile while interacting with a student. They are in the experiment location with the entrance to Wold behind the student and the mostly unseen stairway to the left of the student. This image shows SARAH's size in relation to a person. She is a very good size for a social robot and some students are shorter than her while others are either her height, or taller. Also shown in the image is a small lounge in the right-hand side. Although there are often students in that spot, they tend to go there to work and therefore the noise level is very low. This is ideal for the experiment location. Sometimes people who are studying in that spot grow curious of the robot next to them and decide to participate in the experiment. By looking at the image one can see that there is still a lot of room left between SARAH and the other side of the hallway. This prevents SARAH from becoming an obstacle since traffic can continue to flow naturally when she is present.

3.3 Selecting Experiment Variables

Social robots are capable of having many customizable features including facial expressions, gestures, sounds, eye-gaze, and others. For this experiment I had to choose which features to implement on SARAH, the social robot.Facial expressions allow the social robot to emote while simultaneously speaking or executing other non-verbal cues. Therefore, I decided that SARAH would use facial expressions throughout the entire experiment process. I had also planned to use sounds in combination with nonverbal cues. Sounds in this situation are non-word utterances that are able to convey meaning. Examples of these kind of sounds

include ”uh-huh”, ”mmhmm”, and ”ahhh”. Other sounds that were considered include various robotic beeps that convey certain emotions such as happiness, sadness, and anger. I decided to not include sounds because the addition would have multiplied the amount of experiment groups and reduce the amount of data available for each group. For the same reason, I did not have groups in which facial expressions were not used. The focus of the project became gestures that are analogous to human gestures. Gestures were chosen because they are a very dynamic form of non-verbal cues. Due to SARAH's physical limitations these gestures only use movement from the moving base.

3.4 Gestures

Having chosen gestures as the main focus of the experiment, the next step was to design them. The ideal gestures are those that are easily implemented on the social robot and recognized by those interacting with the robot. The first step of the gesture design process involved recognizing and describing gestures that people use in daily conversation. I recognized four distinct gestures that are frequently used in conversation.

- Agreement - verbal or non-verbal feedback that tells the other person that they are listening or agreeing. Examples include nodding one's head or saying ”yes”, ”mhm”, or similar phrases.
- Disagreement - verbal or non-verbal feedback that tells the other person that they disagree or are answering ”no”. Examples include head shaking or saying ”no”, ”nuh-uh”, or similar phrases.
- Confusion - verbal or non-verbal feedback that tells the other person that they do not understand and need clarification. Examples include appropriate facial expressions, or looking away from the conversation as if thinking.
- Excitement - verbal non-verbal feedback that conveys a sense of celebration or satisfaction. Examples include appropriate facial expressions, hand gestures, or saying ”yay!” or similar phrases.

The next step was to convert these human gestures into gestures that a social robot with certain physical limitations could use. Since SARAH does not have arms, she would not be able to use hand gestures. She also cannot use her head to emote since it does not have the same degree of motion as a human head. However, SARAH is able to move her body in various directions and combinations. The following set of gestures was created with the intention of being analogous to the human gestures above.

- Agreement - SARAH moves forward a few inches, moves backwards a few inches past the starting point, and then returns to the starting point. This gesture simulates a person nodding.
- Disagreement - SARAH turns left 35 degrees, right 70 degrees, and left 35 degrees returning to the starting angle. This gestures simulates a person shaking their head.
- Confusion - SARAH slowly turns left 45 degrees. For a few seconds the body does not move. Faster, she returns to the starting angle. This gesture simulates a person turning away from the conversation for a moment.
- Excitement - SARAH quickly rotates 360 degrees returning to the starting position. This simulates a person exclaiming ”hooray!” and similar phrases while using appropriate hand gestures.

3.5 Experiment GUI

With the gestures for the social robot designed they needed to be implemented. When the gestures were first created and prototyped the joystick controller that controls SARAH's moving base was used. This made sure that the gesture designs were feasible within the moving constraints of the pioneer, the moving base. However, it is not practical for these gestures to be operated manually by the experimenter during the experiment task. There are several reasons for this consideration. The gestures do not occur independently. They occur in addition to speech and also facial expressions. Operating the controller requires both hands which means that executing the gestures and other features in combination would require at least two experimenters. Also, manual gestures would invariable very during execution. It would be very difficult for

the experimenter to exactly replicate the gestures every single time for each participant. In addition, using the controller to manually operate the gestures would prevent the experimenter from writing down observations about the interactions in real time. This would increase the amount of time that the experimenter would have to spend watching the experiment videos after the fact.

The solution to the problems surrounding manual gesture controls was to create a graphical user interface that allows the gestures to be launched semi-automatically. The experimenter still indicates that a gesture needs to be performed by after the process is selected, it executes on its own. The gestures were translated into python scripts that perform the same actions on the moving base every time. The end result was a an experiment GUI with buttons for each of the gestures. In order to make the experiment GUI even more robust the same was done for the facial expressions and dialogue phrases. Since SARAH is not intended to move outside of the gestures, the experiment GUI is the main tool available to the experimenter. Buttons that combine gestures, facial expressions, and verbal dialogue further streamlined the experimenter's job. The original version of the experiment GUI also contained buttons to launch nonverbal sounds. However, these features were later removed from the experiment design.

3.6 Experiment Task - Participant Perspective

The goal for the experiment task is for SARAH to carry a conversation with a person allowing her to accompany speech with the gestures previously outlined. From initial experiments, it was found that when people had an open ended conversation with SARAH the subject of the conversations varied by a large margin. Some people were very inquisitive and asked SARAH many questions about her purpose and challenged her knowledge. Others were more silly and asked SARAH to dance or even jokingly proposed to her. For the purpose of this experiment it is preferred that the conversations between people and SARAH be similar so that the greatest difference between interactions is whether or not SARAH uses gestures while communication. Therefore the conversation between SARAH and the participants is partially scripted.

SARAH's interaction with a participant begins with SARAH saying "hello, can you help me?" as an invitation. At this point SARAH is standing in her strategic location next to a chair that contains a box for completed surveys, part of the final phase of the task. On her tablet is displayed the robot face program with the default face. SARAH does not move while inviting people. Once it is clear that the person wants to interact with SARAH, she says "thank you for helping me" accompanied by the smiling facial expression. Then SARAH describes the task which is as follows:

> Take an index card from the blue pocket. Ask me the questions on the card.

This dialogue, and other dialogues that SARAH delivers, was longer in the first version. However it became clear that there was a noticeable delay between SARAH's mouth moving and the dialogues being produced by the speakers. After some testing it was discovered that the shorter the dialogue phrases, the less of a delay.

On the index cards were three general trivia questions. Most of them are world facts such as the fastest animals of different animal kingdoms or tallest building in the world. For each question the person was expected to read it and SARAH would respond with either a verbal-only response or a verbal response accompanied by a gesture. The cards were created so that each question was a yes/no question. For each index card SARAH responded with "yes", "no", and "I don't know". However the order of those replies was random so that people who participated in the task multiple times did not expect them in a specific order. Certain gestures and facial expressions are always paired with the same verbal replies:

- The verbal response "yes" is accompanied by the agreement gesture and the smiling facial expression.
- The verbal response "no" is accompanied by the disagreeing gesture and the frowning facial expression.
- The verbal response "hmm, I don't know" is accompanied by the confused gesture and the neutral facial expression.
- The verbal response "yay, we completed the task" is accompanied by the excitement gesture and the surprised facial expression.

The responses are not delivered in the same way. While the ”yes” and ”no” responses are delivered immediately after the question has been read, the ”I don’t know” response is more nuanced. It is also followed by either a ”yes” or a ”no” response. This simulates a person who is initially unsure about an answer, and needs a moment to think. The final response, which can be accompanied by the excitement gesture, only occurs after all three questions on the index card have been asked and answered.

After the question and answer portion of the experiment task SARAH has once more request for the participants.

> Please complete a survey inside the yellow pocket. Place the completed, folded up survey in the blue bin next to me.

The purpose of the survey is to take a measure of the participant’s feelings about the social robot. There are three questions on the survey. The first two questions ask the participants to rank statements on a scale of 1 to 5. A score of 1 indicates that they strongly disagree with the statement. Alternatively, a score of 5 on this scale indicates that they strongly agree with the statement. The two statements are as follows:

- The social robot seems intelligent.
- The social robot seems natural.

The final question on the survey asks the participant to list three adjectives to describe the social robot. The questions allow us to interpret the data two different ways with the first two questions leading to a quantitative result and the final one a more qualitative aspect. After the surveys are completed SARAH gives them a final ”thank you for helping me” and the task is complete. The surveys do not collect a name field or any other field that could identify individual participants. They are also collected in an arbitrary order since the box is open and surveys can easily shift position. Therefore at the end of the data collection period there are two separate collections of data, there is the data from the web cam and the data from the surveys. Individuals on the web cam cannot be linked to a corresponding survey and there are many more participants than there are completed surveys since many participants did not complete the task and reach the final portion.

3.7 Experiment Task - Experimenter Perspective

To the participant it appears as if the social robot, SARAH, is acting autonomously. However SARAH is completely controlled by the experimenter. The experimenter is responsible for booting SARAH, setting her up in the experiment location, and closely monitoring SARAH’s interactions with the participants. The experimenter must make important decisions such as when to have SARAH say ”hello, can you please help me?” and how long to wait between the execution of different pieces of dialogue. However, the job of the experimenter is also very scripted due to the nature of the experiment. Once SARAH is engaged with a person there is a very specific order of instructions that the experimenter must follow.

The experimenter’s role begins by turning SARAH on and the various features that form together the social robot. This includes turning on the display and launching the robot face interface. Also, the web cam must be working properly, otherwise the experimenter can not view the experiment environment. SARAH is then set up in the experiment location. SARAH’s movement is controlled with a joystick controller. The controller allows SARAH to be turned left and right and also can adjust the speed of the moving base. Once SARAH is in position the blue and yellow envelopes must be secured on top of her. These envelopes contain the index cards and surveys that the participants will interact with. The final piece is placing a chair next to SARAH on top of which sits a box for collection the completed surveys after an experiment task has been completed.

Once SARAH has been setup in the experiment environment and all of SARAH’s features are working properly, the experimenter returns to the Crochet Lab where SARAH is controlled remotely. A program on the main computer, inside the Crochet Lab, allows the live video feed from the front-facing web cam to be saved to a file. This allows the experimenter to re-watch the interactions and make additional observations. From the live web cam feed the experimenter waits for a person to be in SARAH’s line of site. Timing is crucial in this stage because if SARAH asks for help too early or too late then the person may not be close enough to SARAH when she speaks. There is also an addition setback during this phase of the experiment.

Due to a network complication there is a delay of about two seconds between actions in the real world and what is displayed on the web cam feed. Therefore this discrepancy must be accounted for when the experimenter is choosing when SARAH must ask for help. It is possible that in the future environmental cameras may simplify this problem. Once the experimenter succeeds in having SARAH attract someone's attention it is on to the next part of the experiment, the task.

The experimenter can usually tell when SARAH has grabbed someone's attention because the person will begin to walk towards the social robot. Sometimes it appears necessary to have SARAH repeat the phrase "hello, can you please help me" because it is possible that the person has taken an interest an SARAH but did not hear that phrase the first time. There is sometimes a small delay between attracting a participant and when the task begins. This is because participants are often curious about SARAH and prefer to ask their own questions prior to completing the task. However, as soon as it is clear that participant is ready to proceed to the next step the experimenter has SARAH describe the task of asking questions from the index card in the blue pocket. As the participant read each question individually the experimenter already knows which response to have SARAH perform. The experimenter has to carefully consider the timing of the response due to the aforementioned lags. Usually this leads to the response being selected slightly before the participant finishes asking the question. This leads to SARAH responding after the question is asked making her seem more engaged than if they had been a pause. With the help of the experiment GUI the experimenter can have the verbal responses on their own or in combination with the gestures and appropriate facial expressions. One or the other is selected depending on which experiment group the participant is in.

After the main portion of the experiment task the experimenter has SARAH ask the participants to complete the survey in the yellow pocket and place it in the box on the chair. SARAH also thanks the participants after the interaction. Throughout the process the experiment notes information about the participant in a spreadsheet. Each row in the spreadsheet represents an interaction. Information that is collected includes the number of participants, their gender breakdown, and whether or not they completed all parts of the experiment. Additional observations are made in the "notes" section. The participant is rarely in the experiment location along with SARAH. However, occasional check-ins are made to ensure that there are enough experiment materials to continue running data collection. After that round of data collection is complete, the video data is saved and SARAH is brought back into the Crochet Lab.

4 Results

4.1 Number of Interactions with SARAH

Throughout the data collection period of the experiment the experimenter kept a running total of all the people who appeared in the experiment location along with SARAH. This number does not capture every single person since the web cam could only capture the people in front of SARAH, however this can serve as a good estimate of how many people had the potential to interact with SARAH. The experimenter also kept track of the amount of people who interacted with SARAH directly. This means that they went up to SARAH and had a conversation with her. The smallest group of people are the people who interacted with SARAH directly and completed the experiment task. The following graphs shows the amount of people in each of the groups above split into the times that SARAH used gestures during conversation and those when she did not.

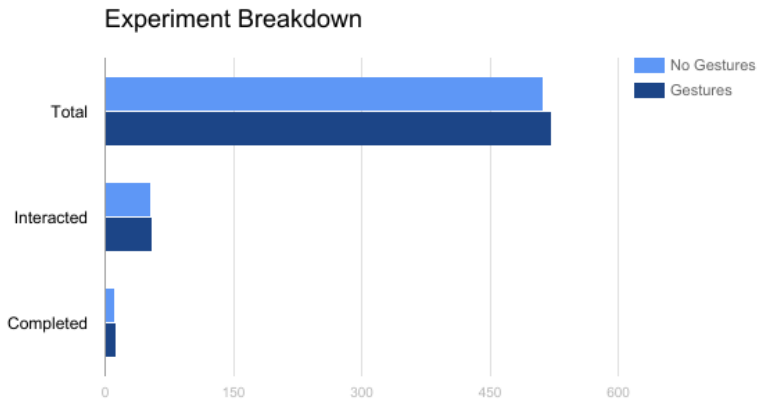

Figure 3: The number of people interacting with SARAH at different levels

Nearly equal amount of data collection time was spent for the two experiment groups (gestures and no gestures). When SARAH used gestures 522 people walked by while when SARAH was not using gestures 513 people walked by. The experiment variable seems to have not played a role in these numbers since the amount of people walking by does not depend on SARAH's actions. Of people walking by during the gestures data collection 55 of them interacted with SARAH which is 10.54%. When looking at the sample when SARAH did not use gestures this becomes 10.53%. There is no significant difference between the amount of people interacting with SARAH in either situation. In the gestures condition 13 people completed the experiment task. This makes up 23.6% of the total peopling who interacted with her in that condition. The complementary value in the other group is 11 people completing the task for 20.4%. This means that more people completed the experiment task when SARAH used gestures. An unpaired t-test was performed on these values to test for statistical significance. To simplify the calculations I treated completing the task as a binary value with 1 meaning that the person completed the task and 0 otherwise. The two-tailed p-value equals 0.6841. The difference between the means of the two groups is 0.03. For there to be 95% or more confidence the value should be between -0.13 and .19. Therefore, there is no statistical significance between the two groups.

4.2 SARAH's Intelligence and Naturalness

The second part of data collection involved participants answering short surveys where the ranked SARAH's intelligence and naturalness. These surveys are suggested by the social robot but completed voluntarily by the participants. There were a total of 24 surveys completed throughout the course of the experiment. The first two questions on the survey ask participants to rank two statements on a scale of 1 to 5 with 1 meaning ”strongly disagree” and 5 meaning ”strongly agree”. The statements are as follows:

- The social robot seems intelligence.
- The social robot seems natural.

Once data collection was complete, the surveys were rounded up and averages were computed to responses to the first two questions respective to the two experiment groups.

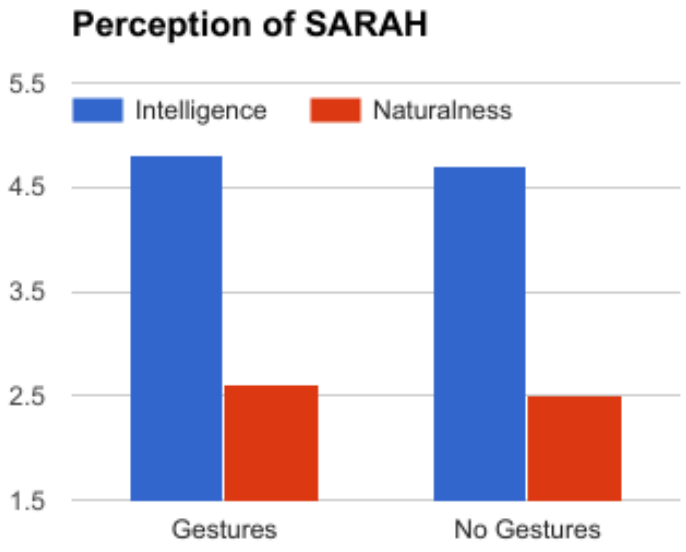

Figure 4: This figure shows the mean rankings for the first two survey questions

The data shown in figure 4 shows that values for both experiment conditions are very close. Another t-tested was calculated in order to ascertain that there is no statistical significance on this set of data. This goes against my hypothesis that SARAH using gestures would receive higher intelligence and naturalness scores from the participants.

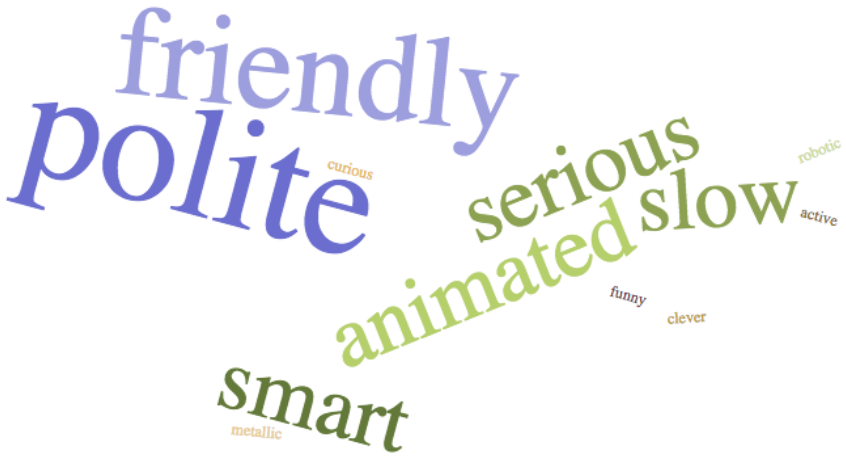

Figure 5: This figure is a word cloud generated from responses to the third survey question.

The final question on the survey asked participants to use their own words to describe SARAH:

Give three adjectives that you would use to describe SARAH, the social robot.

All of the people who completed surveys listed three adjectives for this question. The word cloud shown in figure 5 visualizes the responses given for the final survey question. A lot of people had similar opinions of SARAH noting her politeness and friendliness. However some of the words, such as metallic and mechanical, contrast with those connotations of warmth.

5 Observations

5.1 Data Collection - Early Phase

The term early phase refers to the first week of data collection. The main focus of this phase was to refine the experiment task which included gaining an understand of how much time different parts of the experiment took. This includes everything from turning on SARAH and the various programs that turn her into a functioning social robot to setting her up in the experiment location and also the time it takes to complete each experiment. In this early phase of data collection a lot of time was taken up to set up SARAH. There were some complications that initially took some time to resolve. The first issue is that SARAH's main computer's IP address sometimes changes if the computer restarts. This issue was first observed when various components of SARAH such as the web camera and joystick controller were not able to function after the main launch file was executed. After some debugging the IP address issue was identified and resolved. The issue recurs throughout the period of data collection however as time went by it was more quickly identified and resolved.

Another common issue while setting up SARAH was that after SARAH's main computer is shutdown or rebooted, the tablet is no longer registered as the main display device. When this happens, it prevents the robot face program from running and instead of a dynamic face robot has a black screen. To resolve this issue, a file has to be configured using some terminal commands. SARAH is usually left on which meant that this problem occurred rarely but since the solution is fairly complicated, compared to the IP address issue solution, it tended to cut into the time for data collection. This issue was most persistent during the early phase of data collection. The problem solution was added to a running file of common SARAH troubleshoots and it became less of an issue as time went on.

Aside from identifying and correcting common issues with the experiment setup the early phase was also composed of many interesting observations between SARAH and participants. In the early phase SARAH was a relatively new presence to a lot of people. Even though experiments with SARAH have been conducted in previous years, it appeared that for many people they were experiencing seeing SARAH for the first time. Some participants were more wary than others. While some people acted bashfully and almost fearful around SARAH, others ran to her excitedly eager to see why she was there. During the early phase the experiment GUI had a text field that allowed the experimenter to input any string of text and have SARAH speak it. This made it possible for SARAH to have very open ended back and forth exchanges with people. However, in this early phase many people controlled the conversation asking many questions to SARAH and the experiment task would go incomplete. This led to the decision to remove the text field from the experiment GUI and whenever these open ended questions were asked to SARAH, she would respond with "yes", "no", or "I don't know". SARAH would also insist on moving to the experiment task where she was have more control of the conversation. Although on the surface it may seem counterintuitive to limit conversation between SARAH and participants, this limitation allowed for the conversations to be more uniform and focused on the experiment question around the effect of gestures during conversation.

It was also during the early phase of data collection that it was observed that people around SARAH felt more comfortable in groups than as individuals. There are a few times when someone was interacting with SARAH on their own appearing a bit nervous about the experience. Then a stranger would appear behind the individual encouraging them to participate in SARAH's task. What was most interesting about these type of interactions was that often times the people encouraging the participants had not previously taken part in the experiment themselves. Although they may have interacted with SARAH, or another social robot LINDSEE, in previous experiments there is no way to be certain. In contrast to the positive effect of people in groups, it also appeared that participants in groups were less focused when participating in the experiment task. On average people in groups take more time to complete the task and often try to steer the conversation away from the task or spend a lot of time talking amongst themselves.

5.2 Data Collection - Later Phases

There were some changes in the behavior of people around SARAH towards the later parts of data collection. People who had participated in the experiment before did not participate in the experiment task again. If SARAH invited them they would decline kindly and keep walking. Some people told SARAH

directly that they did not have time to help her out. It is possible that some of the people who declined to participate in this experiment already participated in a different experiment involving SARAH. Other people who had not interacted with SARAH before tended to quit participation around the same phase, SARAH's explanation of the task. When SARAH asks people to take index card from the blue pocket and ask her the questions on it people get a sense of how long the interaction might take. This is a logical time for people to decide if they are willing to go along with SARAH's task.

6 Data Analysis & Conclusions

In the experiment a total of 1,035 people were seen to have been in the experiment location with SARAH. Of those people 10.53% of them stopped to talk to SARAH. This is a very low rate and leads to a small portion of potential data being collected. One way in which the experiment could be improved is to find ways to attract more attention towards SARAH so that there is more data available for the experiment task. More people interacting with SARAH translates to more people completing the experiment task.

There are two significance t-tests that were conducted for the experiment. The first one tested if the experiment group had a significant effect on whether or not people completed the experiment task. This test showed that the connection was not significant. The second one tested if there is a connection between the experiment groups and the intelligence and naturalness rankings given to SARAH by the participants. According to this t-test this was also not statistical significant. There are a few factor that could have lead to these results. The gestures are designed to be very minimal in order to prevent any sense of unease or distress to the participants. It is possible that these gestures were not dramatic or dynamic enough to convey the desired meanings. It is also possible that the participants did not understand the purpose of the gestures or that they were analogous to human gestures.

7 Future Work

In this experiment some of the observations suggest that the gestures may have been unclear. A possible area of research for further work is to focus on the development of gestures and set up an experiment where participants look at a set of gestures and try to identify them. The gestures that are most easily identified would be the most desirable. Gestures in this experiment are limited to body movements do to the physical limitations of SARAH. The use of a different social robot, perhaps one with arms or legs, may open up opportunities for different forms of gestures.

Another idea is to use different non-verbal cues other than gestures in the experiment. Facial expressions could be analyzed with a different robot face program with the ability to express a wider range of emotions. This would allow the researchers to observe people's reactions to different facial expressions and perhaps see which ones are the most effective and recognizable. Related to facial expressions is eye gaze. A social robot use eye-gaze may seem more lifelike than one without. Eye gaze could also be used to refer to objects in the environment.

In future versions of this experiment the experiment task could be different. In this experiment the task was for participants to ask three yes/no questions to SARAH and to have SARAH respond with or without using gestures. This conversation is very scripted and may have played a role in downplaying the scores for SARAH's naturalness. Perhaps the experiment task could only be partially scripted to allow for some degree of free speech between the participants and SARAH.

References

[1] Colleen M. Carpinella et al. "The Robotic Social Attributes Scale (RoSAS): Development and Validation". In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: ACM, 2017, pp. 254–262. ISBN: 978-1-4503-4336-7. DOI: 10.1145/2909824.3020208. URL: http://doi.acm.org/10.1145/2909824.3020208.

[2] Marc Dupont and Pierre-François Marteau. "Gesture Control System for Mobile Robots". In: *Proceedings of the 27th Conference on L'Interaction Homme-Machine*. IHM '15. Toulouse, France: ACM, 2015, 41:1–41:6. ISBN: 978-1-4503-3844-8. DOI: 10.1145/2820619.2825022. URL: http://doi.acm.org/10.1145/2820619.2825022.

[3] Susan R. Fussell et al. "How People Anthropomorphize Robots". In: *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*. HRI '08. Amsterdam, The Netherlands: ACM, 2008, pp. 145–152. ISBN: 978-1-60558-017-3. DOI: 10.1145/1349822.1349842. URL: http://doi.acm.org/10.1145/1349822.1349842.

[4] S. Jain et al. "Analytical Approach for Obstacle Negotiating Capability of Stair Climber". In: *Proceedings of the 2015 Conference on Advances In Robotics*. AIR '15. Goa, India: ACM, 2015, 42:1–42:6. ISBN: 978-1-4503-3356-6. DOI: 10.1145/2783449.2783491. URL: http://doi.acm.org/10.1145/2783449.2783491.

[5] Soyoung Jung et al. "Personality and Facial Expressions in Human-robot Interaction". In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*. HRI '12. Boston, Massachusetts, USA: ACM, 2012, pp. 161–162. ISBN: 978-1-4503-1063-5. DOI: 10.1145/2157689.2157735. URL: http://doi.acm.org/10.1145/2157689.2157735.

[6] Tomoko Koda et al. "From Cartoons to Robots: Facial Regions As Cues to Recognize Emotions". In: *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction*. HRI '10. Osaka, Japan: IEEE Press, 2010, pp. 111–112. ISBN: 978-1-4244-4893-7. URL: http://dl.acm.org/citation.cfm?id=1734454.1734497.

[7] Isabelle M. Menne and Birgit Lugrin. "In the Face of Emotion: A Behavioral Study on Emotions Towards a Robot Using the Facial Action Coding System". In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: ACM, 2017, pp. 205–206. ISBN: 978-1-4503-4885-0. DOI: 10.1145/3029798.3038375. URL: http://doi.acm.org/10.1145/3029798.3038375.