May 2017

Creative Robotics Studio

Mathew Charles Schwartzman
Worcester Polytechnic Institute

Patrick Murphy
Worcester Polytechnic Institute

Steven H. Rangel
Worcester Polytechnic Institute

Turner J. Robbins
Worcester Polytechnic Institute

Creative Robotics Studio

An Interactive Qualifying Project Report submitted to
the faculty of

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the degree of
Bachelor of Science

By:
Patrick Murphy
Steven Rangel
Turner Robbins
Mathew Schwartzman

Project Advisors:
Scott Barton
Craig Putnam

Table of Contents

Tables

Figures

Abstract

The Creative Robotics Studio was initiated as an IQP dedicated to exploring the potential artistic qualities of robotic technologies. This year, we sought to better understand the emotional aspects of human-robot Interaction through studies of robot motion. Specifically, we endeavored to analyze the contexts and kinematic quantities of motion in relation to human affect using a preliminary survey of human opinions on robotic technologies, and an experiment testing human reactions to robotic movements. Using this data, we planned on designing a library of independent motions that can be applied to other robotic designs to facilitate more effective human-robot interactions. Our results, though less conclusive than initially proposed, showed promise for our experimental process and subsequently delivered lexicon.

***Mission statement*:**

The Creative Robotics Studio seeks to explore how robotic motion affects an audience's emotional response towards the robot itself. The team will develop an experiment to empirically measure a spectrum of emotional responses from fear or disgust to comfort or congeniality and then use those findings to develop a set of motion guidelines and tools for increasing the accuracy and efficacy of Human Robot Interactions.

Project Objectives:

1. Develop a library of gestures with the intent of exploring the emotional aspects of Human-Robot Interactions
2. Map the gesture library to the ABB IRB 1600
3. Design an experimental methodology measuring affective responses based on the gesture library.
4. Conduct the experiment and using the results develop guidelines for using the gesture library to improve HRI

Acknowledgments

The 2016-2017 Creative Robotics Studio relied heavily on a tool conceived by the 2015-2016 CRS team, primarily built by team member Graham Held. We met with Graham multiple times throughout the projects terms, during which he helpfully instructed us in the use of his RoboBlender plugin.

We would also like to thank Ryan Mocadlo for acting as Lab Monitor over the weekend during our final experiment. Additionally, we appreciate the entire Washburn Labs staff and their assistance and advice in working with the ABB Robot.

We would also like to thank Ryan Mocadlo for coming to act as lab monitor on the weekend of our experiment, as well as the other Lab Monitors and staff for their assistance in working in the lab and on the ABB IRB 1600.

Of course, we would also like to express our most sincere gratitude for the guidance provided by our two WPI faculty Advisors. Professor Scott Barton provided valuable insight and extensive feedback on our experiment process, as well as our working final paper. Additionally, professor Craig Putnam was crucial in our learning to use the IRB 1600, as well as the other lab resources, and took time out of his weekend to participate in our final experiment. The Creative Robotics Studio project would not have been possible without these phenomenal individuals and their advice.

Executive Summary

Introduction

Industrial robots are generally constructed with no common method of communication. In an industrial setting the movement of robots is purposefully created to be functional; there is no need for an industrial robot to express anything or communicate with a viewer when acting alone. However, when an industrial robot is working together with humans, there is a need for a means of communication. Factory environments are often loud and would make verbal communication impractical. A gesture based communication method would be ideal however no standardized library for movement patterns and meanings exists.

During the course of the project, the team planned to address these issues by creating a gesture library using an ABB IRB 1600. In this lexicon, movements are created and recorded that are designed to elicit specific emotional responses. In order to ensure that these motions are eliciting the appropriate emotion response from the user, an experiment was conducted in which subjects observed the motions from the library and then recorded their affective responses. This data was then analyzed and used to see if the gestures correlate to the intended emotions. Demographic information about the participants was also be recorded in order to determine the population being studied.

Methodology

We developed a library of gestures containing names and information about each one. When organizing our lexicon, it was necessary to define parameters and constraints so that others may use it effectively. For this reason, a Task Space Region (TSR), was determined. A TSR is a constraining representation of the 3-dimensional reference frame in which the proposed gesture takes place (Berenson, Srinivasa, & Kuffner 2011). Typically, when using TSRs to define robot-controlled motion, reference frame w would be centered at the origin of an object to be manipulated by the robot, but as our gestures are not designed to exclusively manipulate objects or environmental elements, some of them are defined by two TSR’s, with the latter defining the final pose (Holladay & Srinivasa, 2016). For example, our lexicon uses one TSR to define the space in which a robot would wave. This is because waving is a simple, repeating gesture that has no starting and ending point. With these TSRs, we developed a lexicon that can be applicable to any robot with the appropriate anatomy.

After developing the lexicon we mapped the gestures to an ABB IRB 1600. Several of the gestures used an HTC Vive VR system which has controllers that are capable of tracking position in 3 dimensions. We used this to track the motion of an actors' hand. In order to convert the motion capture data to a robotic animation, we used the popular animation software Blender with a plugin that allows the user to more easily animate ABB robotic arms. The Blender plugin can calculate an animation given the data. Once these animations are completed, the animation software can output that animation into an array which contains joint positions over time. This array can be converted into RAPID code, which is a high level programming language used to control ABB industrial robots that is similar to C in syntax. An IRC5 Compact controller, which controls the ABB arm, reads this RAPID Code and controls the arm.

Findings

For our experiment we were able to gather 30 participants. However, not all participants gave valid responses, so some responses had to be removed. In particular for questions 1 through 4 there were 27, 26, 26, and 29 responses respectively. We performed the principle components analysis anyway to show how it might be done, and to suggest recommendations for future study.

Before conducting the principle components analysis we ran the Bartlett test of Sphericity over the correlation matrix for the responses to ensure that the samples are not from populations with equal variances. The results, which suggest that for all four questions there are variances between the gestures, can be seen in the table below.

Table 1 Significant Variance as Shown by Bartlett test of Sphericity

Question	Chi Squared	P Value	Degrees of Freedom
Q1 - Elicited Valence	256.2364	2.62E-05	171
Q2 - Elicited Arousal	352.701	1.06E-14	171
Q3 - Percieved Valence	234.4598	0.0009183929	171
Q4 Percieved Arousal	240.5394	0.000362136	171

The principal components analysis for elicited valence responses resulted in the four rotated components accounting for 18%, 17%, 15%, and 10% of total variance respectively. The first component, RC1, had high factor loadings for the hand animated 'Presentation' gesture, and the 'Point' gesture at 50%, 75% and 100% speeds. These gestures all had the robot in an arched posed and were performed at a high speed which suggests that RC1 is related to speed or arched pose. RC2 had high factor loadings on the 'No, 'Taunt', 'Presentation natural motion', and 'Point at 25% Speed' gestures, all of which are low speed gestures. RC3 had high loadings on the 'Bored', 'Cautious', 'Dance natural motion', and 'Look Around natural motion' gestures. Only two gestures had high factor loadings on RC4, which makes it difficult to identify.

The principal components analysis for elicited arousal response resulted in two rotated components which accounted for 30% and 25% of the total variance respectively. All of the gestures that load highly on the first component could be qualitatively described as slow gestures, and all of the gestures that load highly on the second component can be described as fast gestures. The three 'Point' gestures at 50%, 75%, and 100% speed had high loadings on the second component. This leads us to believe that the first factor of arousal is 'melancholy' and the second 'aggressive' or 'high velocity' motions.

The principal components analysis for the perceived valence responses resulted in five rotated components which accounted for 15%, 13%, 13%, 12%, and 12% respectively. Unfortunately, the results are perhaps to muddled to interpret, with only half of the gestures having a high factor loading.

The principal components analysis for the perceived arousal responses resulted in four rotated components which accounted for 17%, 15%, 12%, and 11% of the total variation respectively. The first rotated component, RC1, shows high factor loading for the three high speed 'Point' gestures, which strongly suggests that RC1 is related to high speed. RC2 shows a high loading on 'Bored', 'Desolation', 'Taunt', and 'Wave'. These gestures are all low speed gestures which suggests RC2 is related to low speed. RC3 shows high factor loadings for 'Cautious' and 'natural motion Look Around' both of which have sudden jerky movements, suggesting RC3 is related to smoothness. RC4 shows high loading for 'Curiosity natural motion', 'Dance natural motion', and 'Present'. These three gestures all have a large scale or TSR, which suggests RC4 is related to gesture size.

Responses by Demographics

To determine whether or not there was a significant difference in the responses given we first determined whether the variance was similar for responses by females and males. A T-test was then conducted for each gesture to determine whether or not the samples came from the same underlying population, that is whether or not males and females had different responses. Across all 4 questions and 19 gestures there were only 14 instances where the probabilities of responses being drawn from the same population was below .20, as can be seen in the table below. In fact, the ‘Elicited Arousal’ responses were likely drawn from the same population indicating that males and females both reported similarly when asked how excited the gesture made them feel.

Table 2 Results of T-Test to determine whether or not samples came from same underlying population

	Gesture																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Elicited Valence	0.22	0.73	0.57	0.16	0.20	0.16	0.08	0.36	0.32	0.39	0.02	0.30	0.40	0.38	0.60	0.29	0.52	0.12	0.07
Elicted Arousal	0.47	0.25	0.27	0.93	0.46	0.50	0.98	0.72	0.21	0.62	0.35	0.77	0.43	0.65	0.22	0.98	0.59	0.98	0.61
Percieved Valence	0.35	0.27	0.90	0.75	1.00	0.45	0.86	0.86	0.94	0.19	0.05	0.55	0.92	0.60	0.01	0.22	0.17	0.92	0.65
Percieve Arousal	0.13	0.26	0.14	0.91	0.83	0.47	0.56	0.72	0.72	0.31	0.24	0.90	0.15	0.76	0.46	1.00	0.49	0.03	0.28

Gesture Lexicon

As mentioned previously, a key objective of this project was the development and publishing of a lexicon of movements effective in communicating ideas and emotions to humans. We hope as this library continues to develop, it will prove an invaluable tool in studying and improving HRI.

In classifying gestures, we analyzed their kinematic properties and their common intents, categorizing them in any of the four movement types listed previously. Following the Task-Space Region template, we also found it useful to track movement of the end-effector exclusively. We theorized that this would correlate to interactors' comfort zone and overall confidence in the experiment. The information for each gesture can be found in the gesture lexicon which is attached to this report.

Conclusion and Recommendations

Our findings overall were decidedly less significant than we had hoped. However, with this experiment we hoped to define methodologies to more effectively experiment in creative robotics. That said, a number of interesting correlations were shown between quantities like TSR and comfort zone, as well as those in the preliminary survey results. With smaller sample sizes and limited time to execute our experiment, our conclusions are primarily focused on improving the experiment methodology and data collection methods.

After the preliminary survey and the final experiment, we received a lot of comments and feedback about our execution of certain things, including question phrasing and lab setup. One student asked if we had considered where observers were standing, indicating that each respondent to the questionnaire had a different view of the displayed gesture. This issue was something we had considered but unfortunately did not have time to integrate into the experiment. Another student, upon receiving our debriefing document, voiced their opinion on our hypotheses, stating their unease with higher-velocity and larger-TSR gestures, while identifying smaller, slower gestures as relaxed and even “cute”.

Future Works:

A future experiment could improve upon this experiment in at least three ways. As mentioned before with fifteen subjects standing around the robot not all subjects were facing the front of the robot which was the area the gestures were designed to be viewed from. Another way in which future works could improve upon this experiment is by having more subjects participate in the experiment. The third way in which a future experiment could improve upon this experiment is by have a single method of gesture creation.

The space in which the robot is situated allowed for a maximum of fifteen participants to view the robot at a time. However, this number of participants at once prevented all participants from viewing from the front of the robot. Since the gestures were designed to be viewed from the front some elements of the gesture could have been missed when viewing from the side. Additionally, since the viewpoint of each participant was not tracked, this introduces an unknown variable that could change the validity of results. Ideally a smaller number of people would be brought into the experiment at a time to reduce the problem of location of subjects having different views. This would be possible since we had estimated one-hour time slots and had finished each trial within thirty minutes.

The minimum number of participants for the principle components analysis to have significance was fifty. However, we only had thirty participants. This limited the capability to determine underlying factors. The sign up for the experiment was not announced until within a week of the experiment time. Additionally, the experiment was not at an ideal time. Both of these problems are due to late scheduling of the experiment. By the time we went to schedule the experiment most of the ideal time slots were already taken. As noted above each individual trial could be shorter which might encourage more participants as well as allow for more flexibility in time slot scheduling. Instead of scheduling a three-hour block, multiple shorter blocks could be scheduled. This would increase the chance that potential participates would be available to participate.

As mentioned above some gestures were animated by hand and others used motion capture data from an HTC Vive. The natural motion gestures tended to have more keyframes in development. Additionally, the natural gestures contained the actor’s hand shaking on a minor level. This made the gestures contain a bit of shaking that was not present in the hand animated gestures. At the same time, they more closely simulated how the gesture would be performed by a human. A future work could use the already established code for turning data from an HTC Vive into blender animations as well as the code from the previous group that allowed blender animations to be turned into rapid code to animate gestures more quickly.

We designing the gestures to use for the performance, most gestures were made to mimic a particular gesture of a human. As a result, many of the gestures had underlying context to the particulars of the motion. This resulted in the suspected factors not having proper variations that held the other factors constant. The point gesture was run at several speeds to test if a variation of speed keeping other factors constant would cause a change in responses. One potential way to better represent the suspected factors of trajectory, speed, and acceleration would be to have a set number base gestures that have different trajectories and run each with the same number of variations in speed and acceleration. For instance, if five gestures were to be created then each one would be run with five speeds and five accelerations leading to 125 different animations to run. This would better represent the effective space of the suspected factors.

Authorship

Executive Summary	All; ed. Patrick Murphy
Introduction	Turner Robbins; ed. Steven Rangel
Background	
• **Human Robot Interaction**	Mathew Schwartzman; ed. Turner Robbins
• **Problems in HRI**	Mathew Schwartzman; ed. Turner Robbins
• **Gesture-Based Communication**	Steven Rangel; ed. Mathew Schwartzman
• **Experiments in HRI**	Turner Robbins; ed. Patrick Murphy
Methodology	
• **Generating the Gesture Library**	Mathew Schwartzman; ed. Turner Robbins
• **Mapping to ABB Robots**	Steven Rangel; ed. Patrick Murphy
• **Designing the Survey and Presenting Gestures**	Mathew Schwartzman & Turner Robbins; ed. Turner Robbins
Results	
• **Survey Results**	Mathew Schwartzman; ed. Mathew Schwartzman
• **Experiment Process and Results**	Turner Robbins; ed. all
• **Responses by Demographics**	Turner Robbins; ed. all
• **Gesture Lexicon**	Mathew Schwartzman, ed. Steven Rangel
Conclusion	
• **Conclusion and Recommendations**	Mathew Schwartzman; ed. all
• **Future Works**	Patrick Murphy

Introduction

Creative robots are incorporated into a number of mediums from theatrical performances with human and robot collaborators, to music performed or created by robots, or even kinematic displays that utilize robots. Ignoring philosophical questions about whether or not robots can truly be creators of art, it is important to understand how robots are perceived in a creative context; to understand how the gestures and movements of robots acting in a creative form influence the perception of onlookers. In an industrial setting the movement of robots is purposefully created to be functional; there is no need for an industrial robot to express anything or communicate with a viewer when acting alone. However, in an artistic setting every aspect of what a robot does imparts meaning and must be designed with thought and intention. What robotic motion expresses, whether by intention or accident, is just as important as what robotic motion accomplishes.

As industrial robots, in particular, are generally constructed with no common method of communication, designers resort to non-verbal means of communication such as gesture. In order to design these gestures, one must first study the connections between emotion and expression in humans. Socially Situated Robots, that is, robotic systems for which human interaction plays a key role, can be programmed to exhibit simplified versions of human characteristics, such as natural cues, distinctive personality, and even emotions. Their important roles in research, manufacturing, and education necessitate a set of metrics for quantifying the actions these robots and their human users.

During the course of the project, the team planned to address these issues by creating a gesture library for an ABB IRB 1600. In this lexicon, movements are created and recorded that are designed to elicit specific emotional responses. All of these movements are capable of easily being converted into interpretable commands for a robotic arm. In order to ensure that these motions are eliciting the appropriate emotional response from the user, an experiment was conducted in which subjects observed the gestures and then recorded their affective responses. This data was then analyzed and used to see what features of the gestures caused valence and arousal responses. Demographic information about participants was also recorded in order to understand the population being studied.

Background

Human Robot Interaction

As robotics technology advances, human households and industrial environments further demand human-cooperative intelligently-designed robots. One survey predicts that consumer robotics shipments will double in only 2 years, and then again by 2020 (Wheelock 2017). As the role of robots in day-to-day life grows, effective human-robot cooperation necessitates the implementation of non-verbal communication. Several aspects of human robot interaction including human dynamics, human spontaneity, and human safety need to be addressed. In one model, developed and analyzed by researchers at Tohoku University, it is assumed that a robot interacts with a human user only through a cooperatively manipulated object. For example, a large and heavy table is held by both the robot and a human while the robot responds with an equal and opposite reaction force to assist and maintain pressure on the table (Kosuge & Hirata, 2004). However, it only uses one-way communication (from human to robot). When creating robots that are designed to interact with humans, designers must be made aware of how humans perceive robots and how to effectively communicate a robot's intentions to ensure human safety and efficacy of task completion. If robots are to become ubiquitous they must be able to operate within a social context.

Robots in Social Contexts

Human-Robot Interaction is a social field. With home robotics giant iRobot having sold over 14 million units of their popular Roomba to date, people are rapidly acquainting themselves with how robots can safely and effectively assist in day-to-day life. Though a Roomba is self-monitoring and able to navigate and charge autonomously, owners will commonly be able to watch the robot as it navigates their home. Robots like the Roomba are known as **Socially Situated**: they are surrounded by (e.g., situated in) a social environment that they automatically perceive and react to. Tech companies like Google and Samsung are now also producing home-connected devices like thermostats, refrigerators, and surveillance systems that will likely in the future be able to facilitate complete home automation. These devices are **Socially Embedded** (Terrence, 2003): they are in a social environment and interact with other devices, being partially aware of these interactions. This is in contrast to socially situated robots, which react to human interaction automatically and are not strictly aware of such interactions. A socially embedded unit is one that is considered part of a social structure, versus an individual unit in an external environment (Dautenhahn et al. 2002) For example, a robotic secretary would be socially embedded. Integrating these robots into our typical human social structure necessitates an efficient and implicit communication method.

HRI in Military

Robots are already extremely popular in organized military environments. The attractiveness of a mobile, non-human and therefore dispensable unit is only increased when this unit can effectively operate with only a few operating instructions. These instructions must often be simple and efficiently idiomatic in order to ensure simple interaction between operators and robots (Springer, 2013). For these systems, the assumption is usually made that the operators are not robotics engineers. As such, most of the HRI is comprised of a few instructions from a pilot or mission specialist.

HRI in Medical Care

In medicine robots are currently used in teleoperated scenarios for long-distance or high-precision surgery, enabling dangerous and risky operations to be performed by doctors with little danger to patients. Robot-assisted surgery is among the fastest growing trends in medical care. In 2001, surgeons successfully removed a gallbladder using a teleoperated robotic system in a location in France 7000 kilometers from a patient in New York (Pushkar, 2012). A precise and efficient system that can operate independent of expensive training and time cost is invaluable to every major Medical Center. Once again, these robot systems operate primarily through teleoperation instructions from a single surgeon or specialist. HRI is not as easy/useful to study in these cases. However, some medical centers are looking into using robots for autism therapy. According to the CDC, 1 in 68 children in the US are diagnosed with ASD (Autism Spectrum Disorder). This prominent issue is characterized by abnormal development in social interaction and communication, something which numerous researchers plan to combat with robotic toys like the NAO platform.

HRI in Manufacturing

There are a number of companies manufacturing robots that are used in manufacturing and assembly, materials handling, and welding. According to the International Federation of Robotics, over 253,000 such robots were sold in 2015 Their popularity and widespread acceptance can be primarily attributed to reduced labor costs, increased output rate, and the elimination of dangerous, dirty, or dull tasks from human life. Autonomous self-monitoring material handlers can fashion components while large arm-like robots with many degrees of freedom can assemble these components on an assembly line. Ideally in this factory environment, human involvement is minimal, enabling human workers to focus on more important tasks and only be called in for unit maintenance or upgrades. During this time, units are shut down and repairs are made with little or no active interaction between the robots and humans. Units are reactivated and the system continues with no hassle and little to no more human interaction. This relationship can result in a biased view of robots as cold and unfeeling, and impossible to interact with humans in any normal capacity. One company, Rethink Robotics, developed a manufacturing robot named Baxter that displays its eyes so that it can respond to tasks and humans with expressions and gaze (Knight, 2012). Baxter can be easily programmed to complete manufacturing tasks and responds to incidents such as manufacturing pieces falling. In addition, Baxter is somewhat aware of human movement around it and is designed to be able to sense objects in its path to avoid causing harm to operators. As many humans experience more close interactions with autonomous robots, robot manufacturers design robots with more ergonomic shapes and manipulators. In 2005 a design for a Variable Stiffness Actuator was published by the IEEE. This design used tension on drive-train mechanisms to control the precision and “stiffness” of a robot’s limbs (Bicchi, 2005). With lower stiffness, the robot's movements were more imprecise, but also proved more like natural human motion. These designs ensure safety as well as comfort while interacting with human beings.

HRI in Entertainment

Robots have been a part of human culture at least since the age of the Greeks who had the god Hephaestus and his mechanical servants in their mythos. More recent examples, such as R2D2 from Star Wars and Rosie from The Jetsons, have become both icons and inspiring benchmarks for robotics technology. In a more creative context, robotic displays such as “Please Smile” pose serious questions about how people interpret robots as artists, and explore the theory of intentionality (Nam and Choi, 2012). Industrial robots are also making a splash in today’s popular culture, with popstars such as Lady Gaga using several ABB industrial robots in her Grammys performance to make her keyboard ‘come alive’. These industrial robots where programmed by Andy Robot, developer of a creative tool for Maya that allows users to animate robots (Lalwani, 2016). As robots become more prolific in art and culture it is ever-more important to understand how to make human-robot interactions more effective.

Problems in HRI

Robots that work in these fields miscommunication in three key areas: **Articulation**, **Intentionality**, and **Interpretation** (Lu & Smart, 2011). Articulation problems are defined by the mechanical limitations that robots possess. The articulation of robots depends on the design of the robot. Some robots are only able on or within a plane and others are able to move with multiple degrees of freedom. This reduces the space in which a robot can attempt to both perceive and produce communication. Whether due to a device's physical degrees of freedom or to its inability to mimic social cues like body language or vocal intonation, articulation challenges designers to build controllers and actuators that can use motion to generate subtext in their conversations. The kinematic aspects of typical robotic motion are shown in figure 1 In a measurement of human-vs-robot movements, kinematic quantities of a typical industrial robot such as speed and acceleration were shown to exhibit precisely linear graphs, resulting in what human observers called "unnatural" or "jerky" motion, as shown in figure 2 (Baber et al, 2016). This is in stark contrast to measurements of human movement, which, though less precise, were much smoother and more natural (see figure 2). .

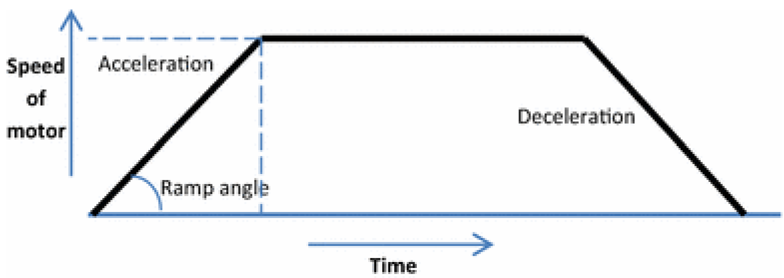


Figure 1 – A speed vs. time plot for 'functional' robotic motion

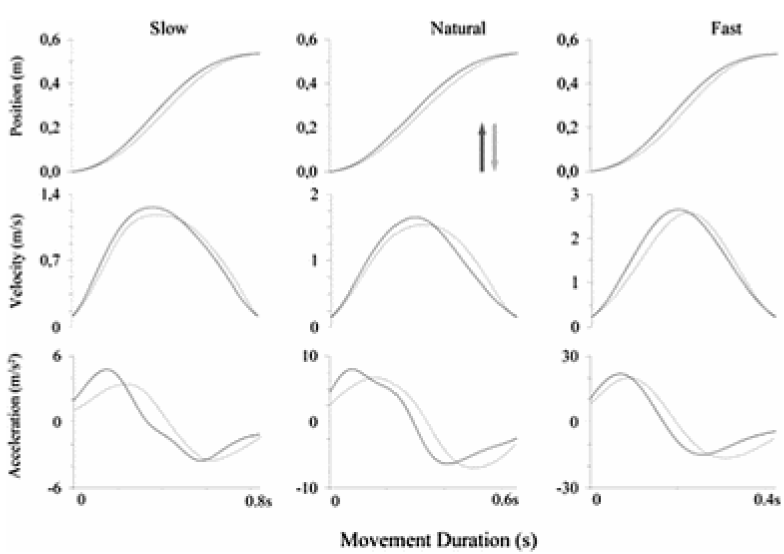


Figure 2 – An example of human motion

Often the intention of robotic motion can be ambiguous to human viewers, which decreases human-robot task effectiveness. For example, if a robot exhibits a simple horizontal sweeping gesture it could be a deictic motion or a metaphoric one depending on factors like velocity and boundary angles. These quantities can be defined by a time series of TSRs, or Task-Space Regions. A Task-Space Region is a definition in 3-dimensional space of a system's boundaries and positions (Berenson et al. 2011). This region assists in creating detailed frameworks for a robot's movements, especially of the end-effector. This can clarify intentions with relatively minimal environmental context. If TSRs are plotted along a time axis, they can define the change in active space that a robot needs.

The last communication issue, Interpretation, arises due to the disconnect between the actions a robot performs and the surrounding context. A modern robot may not be as aware of its environment as a human, and thus cannot make minor adjustments to its state. In the case of an industrial ABB robot, units are designed as simple but precise 6-DOF arms with interchangeable end-effectors, like that shown in figure 3. These robotics arms are quite versatile and strong, but this functionality in movement articulation can lead to interpretation issues.

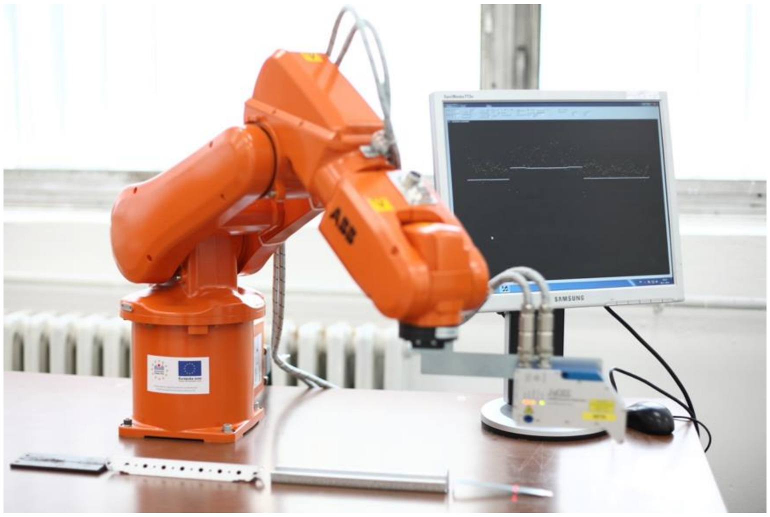


Figure 3 – An ABB robot

Using Gestures and Motion in Communication

One particularly interesting aspect of HRI is motion in relation to communicating and influencing emotion. It is worth noting that some humanoid robots designed for human communication make simple head or arm movements to convey different ideas/context (Li, 2011). These machines are designed specifically with HRI in mind. However, the vast majority of modern robots are built with a focus on practicality and function, such as those built for an industrial environment. In a joint study between the University of British Columbia and General Motors developed new methods of HRI to facilitate close cooperation between humans and robots in an industrial environment. In their experiment, human test subjects were shown video recordings of a human and a robot interacting. At the end of each video, either the human or robot would use a motion to issue a command. The test subjects were then asked what the other should do in response to the command, as well as how easy it was for subjects to understand the gesture. Implicit information like this is often expressed in social interactions through gestures and cues such as body orientation. In human interactions, gently pointing is often used to indicate an object or direction, and is a supplement to vocal communication. In contrast, pointing quickly and sharply can also be used to indicate emotion such as aggression or anger. This shows how different kinematic features can change the interpretation of a motion, and why it is important to take users' affective responses into consideration when designing a solution for increasing the efficacy of human-robot interactions.

Gesture-Based Communication

Many HRI situations call for a more versatile and bilateral model of communication in order to facilitate effective cooperation. A more effective communication model might include both recognizing and reproducing expressive motions or gestures. Several research projects have explored human tracking and movement recognition with the hopes of generating equivalent motions in mechanical systems. In one study by Huang and Mutlu they determined that the best method for studying how people use gestures was through a scenario in which a narrator describes a multi-step procedure. In their example a presenter narrated the process of making paper with a projected visual aid. Some of the more important and exemplary motions that were selected included movements like pointing and size indication. This study then divided motions, hereafter referred to as **gestures,** into 4 distinct groups: Deictic, Iconic, Metaphoric, and Beat gestures. Deictic gestures include object indication, as well as direct references to processes or groups (ex: Narrator points at a particular person or thing). Iconic gestures are used to emphasize action verbs or adjectives (ex: Narrator uses distance between his hands to indicate width of object). Metaphoric gestures are characterized by making abstract motions that commonly refer to cultural language ideas (ex: Narrator sharply rotates his hand to refer to the "next" day). Finally, beat gestures are generally simple and short, presenting the next idea or concept to be explained (I.e. a sharp hand jerk when making a list). These four categories classify one aspect of human motion. Our project features of gesture elicits and influences the perception of affect.

Motion is an important aspect of human communication as can be seen by the number of colloquialisms relating to motion such as “head hung in shame”, “jumping for joy”, and “quaking in his boots”. Many studies have shown that humans can perceive emotions through analyzing motion, and several studies have demonstrated humans can identify emotions used in expressive performances such as dance (Sawada et al. 2003, Camurri et al. 2003, Shikanai et al. 2013). A 2016 study performed by Paul Bremner of the University of the West of England and Ute Leonards of the University of Bristol, tested if humanoid robots performing a gesture along with speech was as effective at communication as the same audio with a video of a human performing the same gesture (Bremner & Leonards, 2016). The study had a robot perform a gesture at the same time as a verbal indication of the gesture being performed and then asked participants to select which picture best depicted the action from multiple images. The same verbal prompt was repeated with a video of a human performing the same gesture. All verbal prompts had two possible actions they could reference and were repeated with gestures for each action. When given both verbal and visual cues, it was found that the robotic gestures were equally likely to be properly identified as the recording of the human performing the same gesture. The study had one gesture where the robotic performance was significantly lower than the human performance. This was thought to be caused by lack of articulation in the robot’s hand.

Lack of articulation, which is often influenced by the number of degrees of freedom a robot possesses, is only one way in which robots vary from human form. The more human a robot appears, the more likely humans are to empathize with it, until it reaches the uncanny valley, as seen in figure 4.

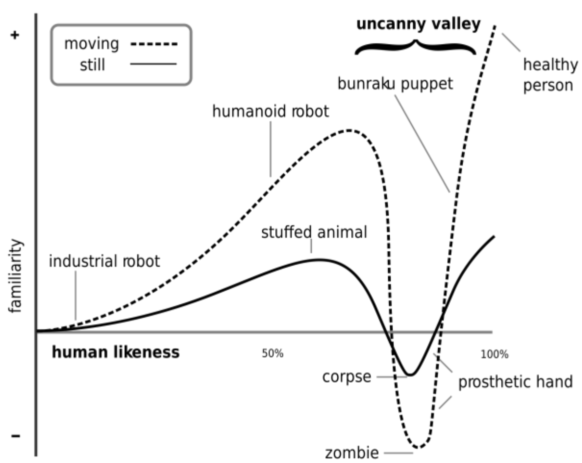


Figure 4 – A graph depicting the drop in familiarity known as The Uncanny Valley

The Uncanny Valley as a measure of comfort indicates a direct relationship between the visual similarity of robots to humans and subjects' comfort with those robots. This may indicate that robots that have gestures and gesture features similar to humans may be more comfortable to interact with, and may increase interpretability of cooperative interaction. In contrast, in animation and film, a skeleton's gestures are commonly over-exaggerated to compensate for unrealistic figure construction or modeling (Kopp & Wachsmuth, 2000). When conveying messages and themes to viewers, these exaggerations are overlooked along with these models. This can be observed as an alternative method of dealing with the uncanny valley mentioned earlier. In order to avoid the models residing in the repulsive region of human likeness, animators make the models move in a way that decreases the motion's likeness to that of human motion, which avoids the Uncanny Valley.

Equally important to the motion aspects of a gesture is the robotic embodiment of the gesture, that is to say the form of the robot. We analyzed human and animal versions of several gestures in a natural context and mapped those gestures to an industrial arm while attempting to retain the kinematic and spatial elements of these gestures.

Experiments in HRI

When studying HRI it is important to consider these social and emotional effects that robots have when interacting with humans. This is compounded by the significance of social interaction to human well-being and communication. Take for example, the aforementioned issue of intentionality. When attempting to collaborate on tasks, subjects said that functional robot motion (motion generated without attempting to be predictable or legible) made it difficult to tell what the robot was trying to do. This made the subjects less trusting of the robots and decreased task efficiency (Dragan et al. 2015). When robots, even industrial ones, are programmed without giving thought to how the motion affects the human emotionally, it can produce motion with ambiguous intentions. This makes human users less trusting and slower to collaborate (ibid). In order to understand how robot motion affects human perception, it is important to know what work has been done around the measurement of emotion, the human perception of gesture, and human perception of robot gesture.

Measuring Emotion

Emotions can be described as physiological and psychological reactions to events found to be significant to the perceiver. Developing a model to represent the emotional space has been a significant focus of psychologists. For example, one such model posits that all emotions exist in a two dimensional space defined by the arousal and valence axes as can be seen in figure 5 (Russel 1980).

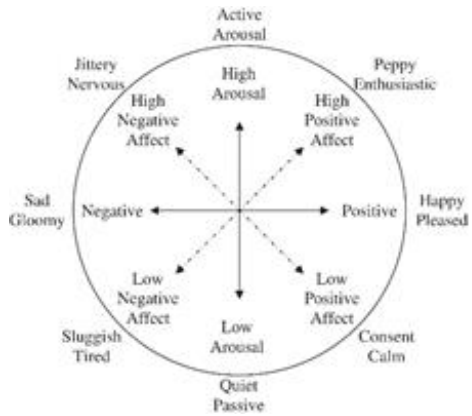


Figure 5 - The Circumplex Model of Affect

Increasing evidence has suggested that there is likely a third dimension to affect, but there is some disagreement on how it affects the model of emotion or how it is labeled (Russel and Mehrabian 1977, Daly et al. 1983). Daly suggested that intensity is the third state, and it rises from the Circumplex Model of Affect to form a conical model as seen in figure 6. In this model more intense emotions would lie at the base of the cone, and more neutral emotions would lie at the tip.

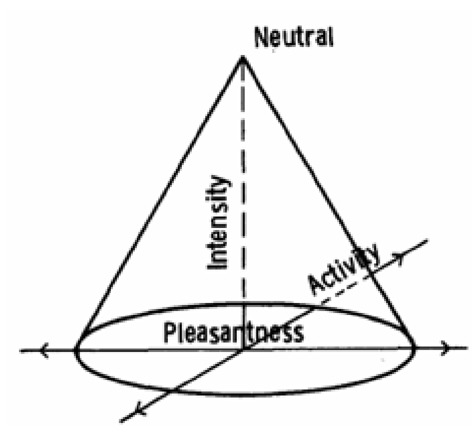


Figure 6 - A representation of the Circumplex Model of Affect with a third dimension, intensity

Finally, the Pleasure-Arousal-Dominance (PAD) model suggests that the third dimension, dominance, is the amount of control over a situation experienced (Mehrabian 1996). It is important to understand that these are all models for the affective space, and each one seeks to explain some degree of variance in that model with the number of dimensions and the representation of those dimensions in a multi-dimensional space. Both the Pleasure-Arousal-Dominance and Circumplex model are widely accepted and have been used for numerous studies to measure affective responses.

In 2003, an experiment conducted using the Kismet robot (seen in figure 7) sought to explore how a robot using emotional models derived from ethological observations could effectively communicate its desires in a social interaction with a human subject. The Kismet robot perceives the users' affective state by using auditory and visual sensors to interpret facial and vocal cues. In addition, the Kismet robot can express itself through the use of facial expressions, body posture, gaze, and 'vocal babbles'. The researchers modeled Kismets emotional space using a 3D model of affect, and used its drives to influence how it moves through that space. The experiment found that users were able to effectively identify Kismet's facial expressions when viewing both still images and video capture of the Kismet robot. In addition, when asked to display affective intent towards Kismet (approval, attention, prohibition, and soothing), the users used Kismets social cues to determine when their intent had been effectively communicated. Users also displayed affective mirroring; when Kismet lowered its head and closed its body language users mirrored the affective cues (Breazeal 2003).

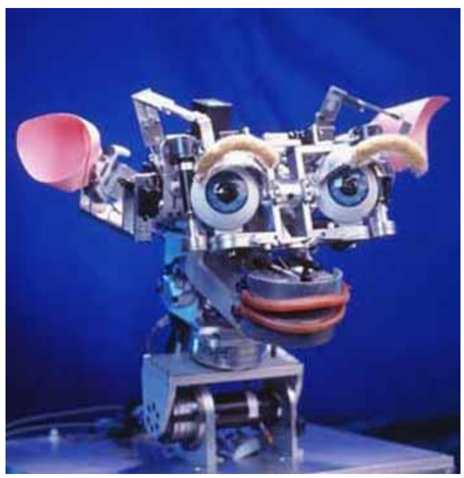


Figure 7 - The Kismet Robot

Perceiving Motion

Human gesture is socially expressive and emotionally dense. Take for example, the art of dance, in which certain choreographed events and motions can be used to affect an audience's emotional response, specifically that arousal is more easily manipulated than valence (Stevens et al. 2009). In a 2009 study researchers seeking to develop a means for continuously measuring emotion used PDAs to record two dimensional emotion data taken in response to dance performances. The study found that choreographed events could be used to easily manipulate an audience's arousal response. Although it's important to consider that these performances were multi-modal, containing musical performances as well as dance, there can be no doubt that motion and gesture have a social significance and emotional impact (Stevens et al. 2009). One study published in 2012 asked users to select items out of a box and place them in a kitchen area as instructed by a HONDA robot, which can be seen in figure 8. Subjects were either given unimodal (speech only) or multimodal (speech and gesture) commands by the robot. Some of the multimodal commands included gestures that did not match the vocal command. When asked to answer questions relating to their perception of the robots, subjects indicated that the multimodal commands made the robot seem more ‘lively’, ‘fun-loving’, ‘communicative’, ‘active’, ‘engaged’, and ‘sympathetic’. While this research only used representational gestures (deictic, iconic, and pantomimic gestures) it still suggests that human-robot interaction can be greatly improved by giving thought to how gesture influences subjects’ perception of the robot (Salem et al. 2012).

Figure 8 - A Honda Asiimo robot giving multi-modal instructions to a participant

Robot Motion

In an experiment done by Paul Bremner and Ute Leonards from the University of The West of England, they tested whether or not gesture would help people understand robot intentions in an intuitive and efficient way (Bremner & Leonards 2016). This theory was tested using a NAO humanoid robot platform from Aldebaran Robotics. This robot can mimic the entire human body allowing for full body gestures. Their experiments had positive results with a robotic presenter improving the HRI. A humanoid robot is not necessarily required for testing the emotional response of certain gestures. Researchers at the University of British Columbia have performed gesture experiments using a 6-DOF arm with an anthropomorphic end effector as seen in figure 9. They manually animated the arm using human gestures as reference. During the experiment participants were able to actually predict and identify the

Figure 9 An arm with six degrees of freedom indicating objects on a table

gestures of the arm for a given situation (Gleeson et al, 2013). This research was conducted with the intention of developing a gestural lexicon to aid in human-robot collaboration in manufacturing.

Certain aspects of motion, such as velocity and acceleration, can also influence how a motion is perceived; changes in both articulation and interpretation can be linked to kinematic features. For example, a gesture performed with high velocity and acceleration will often indicate a state of high arousal, whereas low velocity and acceleration will indicate low arousal; this was found to hold true in an experiment using a robotic arm to display several different motions, none of which were designed with emotional intent (Sial et al. 2016). A study at the Tokyo Institute of Technology had participants watch dances that were meant to convey certain emotions including joy, sadness, and anger. The participants in this study were able to perceive the intended emotion from the dance, but with different accuracy depending on the emotion they were supposed to detect (Sawada et al. 2003). The various dances were tracked by computers to measure the kinematics involved. The study found clear differences in the velocity of motions from dances that were meant to convey different emotions. For example, dances which participants perceived as angry were made up of motions with more velocity and acceleration than dances that produced different emotional responses. Although dances that conveyed joy and sadness didn’t show a significant difference in velocity or acceleration, “a longer traveled distance was contributed to the joy expression”. Another study used a Roomba and an iCat robot to find how perception of emotion was affected by types of motion. They found that acceleration did not affect valance but there was a correlation between acceleration and arousal. In addition, curvature had an influence on valence, arousal, and dominance, but not all were significant (Saerbeck and Bartneck 2010). This strongly suggests that other kinematic features of gesture could have an impact on movement in the emotional space, and that gestures performed by robots may have a similar (if not identical) impact on the viewer’s perception of affect. This does not confirm if robotic motion can change a viewer’s affective state; however, many studies have been done that show inanimate things such as color, scenery, and other visual stimuli can influence affective state.

Our project compiles many of these movements into a library of information like kinematics, prior research, context variables, and measurement tools to provide future researchers and designers with a baseline for non-verbal human-robot interaction. This endeavor is important because there is currently no standardized method of generating/measuring human-interactive robot motion.

Methodology

Our project seeks to understand how certain types of robotic motion affect human-robotic interactions, specifically in relation to emotional responses. We developed three primary goals: 1) Develop a gesture library that covers a large emotional space. 2) Map these gestures to the ABB IRB 1600, an industrial robotic arm with six degrees of freedom. 3) Present the gestures to human subjects so that the subjects' emotional responses can be recorded. Analysis of the data was intended to allow a mapping of certain emotional and logical responses to the gestures giving us a better understanding of what emotions humans are likely to respond to types of robotic motion with.

Generating the Gesture Library

With the design of intuitive actions in mind, we needed a methodical way to measure and classify organic gestures and quantify their kinematic variables. Previously, we mentioned a study in which four categories of gestures were commonly used in the expression of ideas or the description of a narrative. As these categories were analyzed and compiled with human subjects, animal gestures may fall out of these four categories. In addition, these categories did not pertain to emotion so they may not be suited for classifying the emotional content of gestures.

When organizing our lexicon, it was necessary to define parameters and constraints so that others are able use it effectively. For this reason a Task Space Region (TSR) was determined. A TSR is a constraining representation of the 3-dimensional reference frame in which the proposed gesture takes place (Berenson et al. 2011). TSRs are made up of three components: T_w^O, T_e^w,B^w. The first of these values, T_w^O, represents the transform from the origin to the task reference frame w. T_e^w quantifies the offset of the end-effector in the coordinates of reference frame w. The last quantity, B^w as seen in figure 10, is a 6x2 matrix defining the boundaries of the task/gesture in frame w.

$$\mathbf{B}^w = \begin{bmatrix} x_{min} & x_{max} \\ y_{min} & y_{max} \\ z_{min} & z_{max} \\ \psi_{min} & \psi_{max} \\ \theta_{min} & \theta_{max} \\ \phi_{min} & \phi_{max} \end{bmatrix}$$

Figure 10

Typically, when using TSRs to define robot-controlled motion, reference frame w would be centered at the origin of an object to be manipulated by the robot, but as our gestures are not designed to exclusively manipulate objects or environmental elements some of them were defined with an origin at the base of the robot (Holladay & Srinivasa, 2016). For example, our lexicon uses the TSR to define the space in which a robot would wave. This is because waving is a simple, repeating gesture that has no starting and ending point. With these TSRs, we developed a lexicon that can be applicable to any robot with the appropriate anatomy.

Mapping to ABB Robots

In order to properly simulate understandable gestures, we first had to define a method of mapping natural movement to a robotic arm. This presented multiple challenges since the gesture library itself only outlined very abstract gestures. There are several methods that could have been used to generate a path for the robot to follow. Points could be manually plotted using any tool available for the robotic platform to create the gesture. Another method is analysis of existing video of gestures. This can be done in by plotting points on the video frame-by-frame which would reduce the risk of personal bias (when compared to animating the gestures by hand). The downside to this technique is that video data contains two dimensions and cannot easily capture three-dimensional motion. A third alternative is using motion capture data to build a three-dimensional model of the gesture. From this data joint angles, velocities and accelerations could be calculated in a procedural manner to map to the robot.

In order to map gestures to a robotic arm

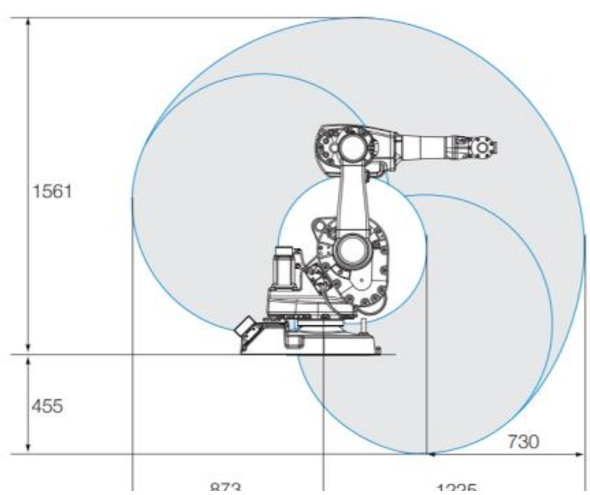


Figure 11

using analysis of existing videos, we originally used a program called Physmo to track motion in videos of the desired gestures. Physmo outputs two dimensional coordinates of points set to be tracked for every frame of the video. Using this data, velocity and acceleration can be calculated and used to form the gesture. For gestures that mostly take place in a two dimensional plane such as the salutation gesture this technique will work, but this method is limited by the fact there is no depth in two dimensional video. One workaround for this limitation would be to capture a video of the gesture from two perpendicular directions and use that data to interpolate the motion in three dimensions. Unfortunately, Physmo ended up not being used due to difficulties obtaining good data from two separate video feeds.

For more complicated movements we attempted to use a visual fiducial system called AprilTags. These tags and the associated algorithm are designed to be able to determine 6DOF position when viewed with a camera (Olson 2011). The first iteration of using this system included on tag on the shoulder, one on the elbow, and one on the wrist of the person to record. A python script was developed to convert the coordinates of the three points into quaternions for the arm to follow. This lead us to discover that mapping of a humanoid arm to a robotic arm with a three dimensional data set was not as simple as in two dimensions. The robotic arm has six degrees of freedom which does allow it to reach any pose within mechanical limits (see figure 11). However, this does not allow for a one-to-one mapping from human joints to robotic joints. The quaternions that were calculated would have required several of the single degree of freedom links in the arm to turn in multiple directions.

Next we decided to only track position and orientation of the end effector. This involved removing the tracker on the elbow and only using the shoulder tracker for a starting position. The problem of converting a pose from a human arm to a robotic arm then became a matter of determining what the zero configuration of the robotic arm should be for our purpose. We were able to get points plotted in Blender, the software we used to animate the arm, but it did not look close to what was desired.

Another option that was considered was using a motion capture system. These systems rely on a series of two or more cameras with fixed orientations. Markers are placed on the subject to track key points that are recorded by the system. This technique was used in a 2016 study “From Human Motion Capture to Industrial Robot Imitation” (Laguillaumie et al. 2016). In this study they used seven points on the human arm to track the motion of the shoulder, arm and forearm. The respective joint angles were run through a script to limit motions to within the bounds of the arm and avoid singularities. We were unable to get access to a standard motion capture system, however we did have access to an HTC Vive VR system which has controllers that are capable of tracking position in 3 dimensions. We used this to track the motion of an actor's hand. Then we fed the data into an inverse kinematic algorithm that output joint angles for the robotic arm to follow the path.

At the same time we manually animated gestures based on existing video footage. The animator would watch a video of the gesture and try to get the robot make a similar motion.

We decided to keep both the motion captured gestures and hand animated gestures to see if there was a difference in perception between the two types of gestures. The motion captured gestures had oscillations that made them seem more natural than the precise motions of the animated gestures.

Once motion data for the gestures was computed the next challenge was to display those gestures on an ABB IRB 1600. In order to convert the motion capture data to a robotic animation, we used the popular animation software Blender with a plugin that allows the user to more easily animate ABB robotic arms. The motion capture data was converted into joint angles per frame. The Blender plugin can calculate an animation given these joint angles. Once these animations were completed, the animation software converted that animation into an array which contains joint positions over time. This array was converted into RAPID code, which is a high level programming language used to control ABB industrial robots that is similar to C in syntax. An IRC5 Compact controller, which controls the ABB arm, read this RAPID Code and controlled the arm. The RAPID Code can also be simulated in RobotStudio, which is proprietary software made by ABB. We used RobotStudio to accurately simulate the arm's gestures before loading them onto the physical IRB 1600. This was done not only to save time but to ensure the arm behaved properly, for example, did not collide with anything.

Designing the Survey and Presenting Gestures

After we assimilated our library of gestures we performed an experiment to gather the participant's affective responses to the gestures and used the collected data to determine which aspects of the motion influence subjects' reactions. In researching how to conduct our experiment, we intended to find best practices for experimental design, to discover how to collect data about the population before the experiment, and how to analyze our data in a meaningful way.

The initial step in designing our experiment was to select the population we wanted to draw subjects from. While we would have liked to gather data about the general US population, it would have been impractical to try and test subjects from across the country so instead we set our eyes on a smaller population: faculty and students from local colleges. We chose not to draw subjects from only WPI because it has a large proportion of S.T.E.M students. To get a baseline for the populations' feelings about robotics and to gather other demographic information we sent out an exploratory survey to the nearby colleges. By reaching 200 individuals from all colleges in the area we could survey the Worcester population of 36,000 college students with a maximum 9% margin of error and 99% confidence level. (Smith, 2013).

When developing the survey, we decided that the most relevant and useful information would be demographics about sex, level of education, age, familiarity with computerized devices, and comfort with different scenarios involving human-robot-interaction. We used a series of categorical questions to determine basic background information; categorical questions are easy for the user to answer and are better for analysis than open ended questions because they require no clean-up. To infer relations between comfort with robotic situations and level of technical skill we will ask participants to rate their comfort with several human-robot interactions on a 5 point Likert-scale. When designing the survey questions, these Likert scales assisted in keeping data discrete and easy to process. The Likert scale uses five points, from zero to four, to indicate an opinion between complete discomfort and complete comfort with provided scenarios. Many researchers regard seven points as the upper-level of Likert Scale efficiency and response coverage, but to save time and ensure compatibility with small mobile devices we provide five points that where numbered with only the extremes labeled (Allen & Seaman, 2007). The midpoint was included to accommodate the potentially significant portion of our sample size which has not formed an opinion on robotics in general (Weems & Onwuegbuzie, 2001). We put several iterations of the survey through a small (15-20 person) pilot study of friends, family, and WPI students. With their feedback, we made small edits to questions and input methods.

After gathering information about the population, we conducted an experiment in which we displayed gestures from the gesture library and asked participants to answer four questions about their own affect as elicited by the robot and their perception of the robot's affect. The first question asked the participant to rate their pleasure (dislike-like) as caused by the gesture on a 9 point Likert-scale with endpoints labeled as displeasure and pleasure, respectively. This was done to measure the participant's valence. We chose to use 9 points because any additional points after 9 fail to add significant accuracy to responses. The second question asked the user to rate their excitement on a 9 point Likert-scale with the endpoints labeled as calm and excited which gave us a measure of the participant's arousal. The third and fourth questions used the same format but asked the participant how they thought the robot was feeling based on the gesture it performed. Each gesture was performed with no end effector.

After the experiment was conducted, we analyzed the data to see which elements of the gestures were causing variation among the affective responses of the participants. We wanted to discover how the kinematic and contextual elements of the gestures influenced the level of emotional impact they had on participants. The simplest way to analyze our gesture data would have been to plot certain aspects of the gestures' motion (velocity, acceleration, jerk, Quantity of Motion, TSR size, etc...) against the participant's affective responses and analyze how valence and arousal are influenced by those variables. Other analytical methods such as Pearson's correlation could have told us how different variables are linearly related to valence and arousal. Finally, we used a principle components analysis to determine which gestures are grouped together and then found what kinematic and contextual elements they share to determine how those elements influence participant's affective responses.

We did by using the gestures as 'variables' and then performing an R-type principle components analysis to discover the elements of the gestures that influence affect. This means we needed at least 5 variables (gestures) for each expected factor, and at least 5 observations per variable, with at least 50 observations. After gathering the data, we chose to measure the sampling adequacy or perform a Bartlett test of sphericity over the data. This told us whether the gestures were correlated enough to have some sort of underlying structure. We then performed the factor analysis and determined the number of factors to extract using a scree plot. Once we determined the number of underlying factors and determined what elements of motion load off of them we began to develop guidelines for motion based off of our lexicon. These guidelines are suggestions for robotic movement in human-robot interactions that can help to develop interactions with more depth and subtext.

Findings

Among the data received from the preliminary survey and experiment questionnaire, we also learned much about measuring effective space and classifying motions. We decided to use three key factors to classify our final gestures: velocity, acceleration, and task-space region. With hand-animated gestures, we controlled gesture speed and trajectory in accordance with the gesture's desired response. When classifying motion-captured gestures, we grouped them similarly, taking values like mean acceleration, peak velocities, and area-of-effect. The results of our experiment can be seen below.

Survey Results

As previously stated, the preliminary survey was tested using a pilot study of 20 friends and family members of the survey designer. Subsequently, the aforementioned improvements were made to the questions of the survey and the survey was distributed to the various universities of Worcester, MA. Including those of WPI, a total count of 87 respondents contributed to the final survey data.

Our goal was to obtain a significant portion of the Worcester College population in order to represent the population of our peers as well as their respective faculty members. Among the schools and departments contacted were Becker's Psychology Department, upon IRB approval. As it happened, 90.6% of respondents were students, all undergrad, from schools like Holy Cross, Becker, NYU, and Texas A&M. The breakdown of Majors can be seen below.

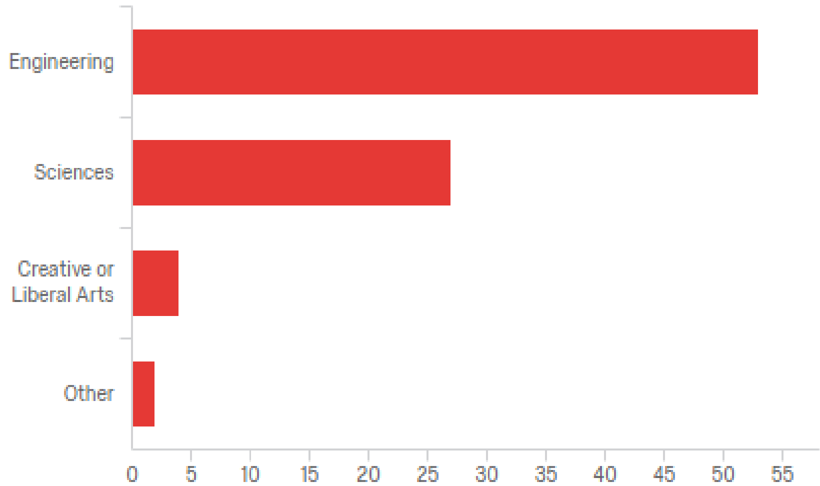


Figure 12 Breakdown of Participant's Majors

As shown, 60% of student respondents major in an engineering field, with the three largest categories being 19% in Robotics Engineering, 10% in Mechanical Engineering, and 11% in Chemical Engineering. Other engineering majors included Biomedical Engineering (5%), Civil Engineering (2%), and Electrical and Computer Engineering (7%). Among the student respondents, 31% majored in science majors, including Computer Science, Biology, Chemistry, and Political Science.

With the other demographic data, we generated cross-tabulations demonstrating some interesting relationships between. When asked about their general interest in robotics, the majority of respondents selected "somewhat interested" or "very interested". We theorized that their experience with Robots would likely correlate to their comfort in situations involving robots, such as in robotic performances, robotically cooked meals, and robotic home security. The graph below breaks down these situations and displays the comfort of those most interested in robotics.

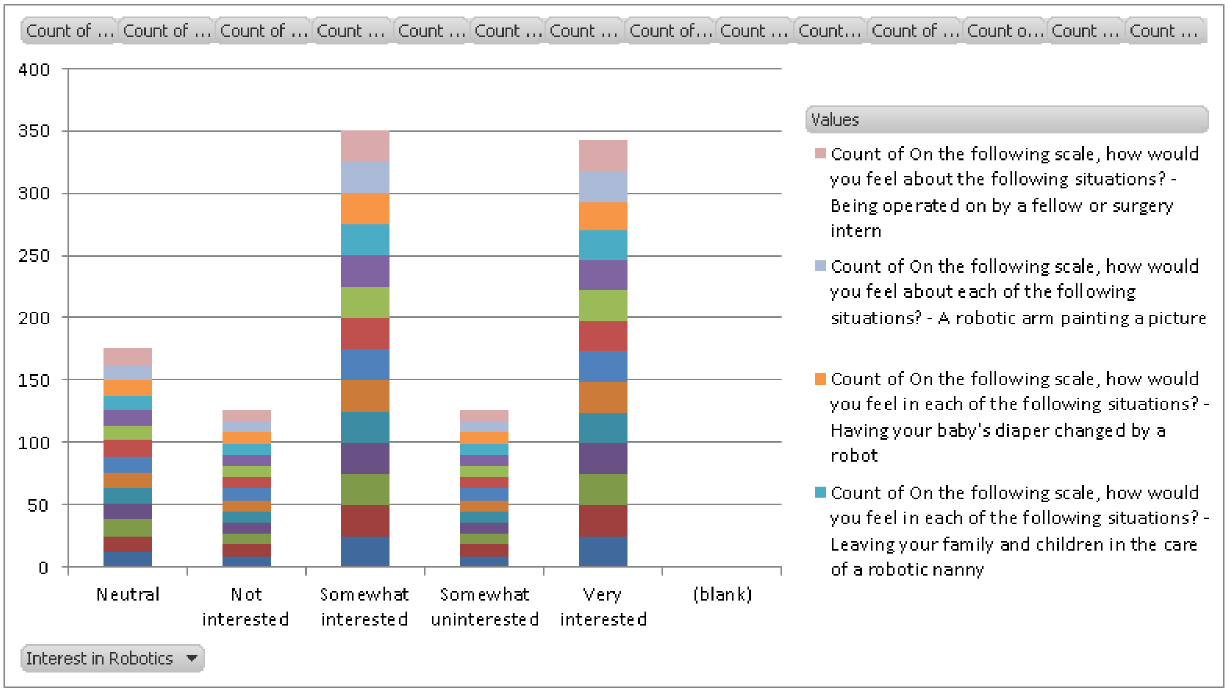


Figure 13 - Comfort Levels for each situation

Here, each colored block shows the total count of comfort level (a sum of Likert responses on a 1-4 scale) for each situation, a quantity shown clearly to correlate to interest in robotics.

Unfortunately, the standard deviation for our subjective Valence and Activation questions was too severe to draw any accurate conclusions, and our sample size of 87 was far too small compared to our desired quantity of 200.

Experiment Process and Results

The number of participants that was required for the factor analysis was at least fifty. Thirty participants took part in the experiment. This resulted in a reduction in the confidence value of the findings. Issues in scheduling the experiment lead to the announcement being only five days before the experiment.

There were two CNC machines around the robot as well as other industrial machines nearby. Further the environment had large yellow mirrors surrounding the robot as part of a safety mechanism. This environment is likely unfamiliar to many of the subjects and may have influenced their responses.

When the participants entered the room for the experiment they chose a position to stand. A semicircle was formed resulting in a different view for each person. Since most of the gestures were meant to be viewed from the front this also could have influenced their responses.

Principle Components Analysis

Our initial goal when developing the gesture lexicon was to show how a set of candid gestures might be mapped to an industrial robot, and then determine what underlying factors influenced subjects' perceived and elicited affect (that is how they thought the robot was feeling, and how they felt themselves) when viewing gestures performed by the industrial robot. This would enable us to provide standards and suggestions to increase the efficacy of human-robot-interaction. The gesture lexicon we initially developed contained 20 gestures and a set of features for each gesture including a description, video sample, and classification of the gesture as deictic, iconic, metaphoric, or beat gesture. As we began to develop methods for transposing those gestures from their original model to a 6 DOF industrial arm, we added implementation specific fields to those gestures to try and describe the gesture in the context of the industrial arm. These fields include velocity, acceleration, end-of-arm tooling, and other contextual features that place our gestures in a hypothetical 'gesture space', and allow us to add more gestures to broadly cover the aforementioned 'gesture space'. This would allow us to better identify what features influenced subjects' affective responses to gestures by performing a factor analysis with the gestures as variables. This helps us understand the underlying factors that influence a subjects' affective response. Unfortunately, due to time constraints and issues with developing a reliable method to turn naturally generated motion into RAPID code for the ABB 1600 industrial arm we could only implement 19 gestures on the industrial arm within the time frame of our project. These gestures were a combination of natural motion gestures and hand animated gestures. 12 of the gestures were generated by animating a scale model of the industrial arm in Blender and then exporting RAPID code of the animation. The other 7 were generated by using an HTC Vive to track an actor's hand or head movements, and then using inverse kinematics to have the robot track the actors hand or head in its own space. The final implementation of the 19 gestures for the experiment used no end-of-arm tooling, and only one kinematic feature of the motion was fully explored. To determine whether or not velocity had a significant impact on subjects' affective responses we displayed one gesture 4 times at varying speeds throughout the experiment. We would like to have implemented gestures with other varying features so that we could more easily see how subjects' affective responses were influenced by things such as acceleration, smoothness of trajectory, or the scale of the gesture (TSR). In addition, factor analysis is usually performed with a great number of responses, as this helps to increase the validity of the results. Normally it is suggested to have $n > 50$ responses, whereas some suggest having $n > 100$. At $n > 50$ factor loadings of .75 have a significance of .95% (.05% chance of the null hypothesis being true). For our experiment we were able to gather 30 participants. However, not all participants gave valid responses, so some responses had to be removed. In particular, for questions 1 through 4 there were 27, 26, 26, and 29 responses respectively. We performed the principle components analysis anyway to show how it might be done, and to suggest recommendations for future study.

Before conducting the principle components analysis, we ran the Bartlett test of Sphericity over the correlation matrix for the responses to ensure that the samples are not from populations with equal variances. To perform this test, we used the 'cortest.bartlett' function in the 'psych' package for R x64. The results, which suggest that for all four questions there are variances between the gestures, can be seen in the table below.

Table 3 Significant Variance as Shown by Bartlett test of Sphericity

Question	Chi Squared	P Value	Degrees of Freedom
Q1 - Elicited Valence	256.2364	2.62E-05	171
Q2 - Elicited Arousal	352.701	1.06E-14	171
Q3 - Percieved Valence	234.4598	0.0009183929	171
Q4 Percieved Arousal	240.5394	0.000362136	171

Next we used the 'princomp' and 'screeplot' functions to determine the number of factors to extract for each question. As can be seen in the scree plots for each graph (see Appendix A), there is a clear elbow at 4, 2, 5, and 4 components for questions 1 through 4, respectively. Finally, principal components analysis was performed using 'principal' function from the 'psych' package with the correlation matrix for the given question, the suggested number of factors, and the varimax orthogonal rotation.

The principal components analysis for elicited valence responses resulted in the four rotated components accounting for 18%, 17%, 15%, and 10% of total variance respectively. The first component, RC1, had high factor loadings for the hand animated 'Presentation' gesture, and the 'Point' gesture at 50%, 75% and 100% speeds. These gestures all had the robot in an arched posed and were performed at a high speed which suggests that RC1 is related to speed or arched pose. RC2 had high factor loadings on the 'No, 'Taunt', 'Presentation natural motion', and 'Point at 25% Speed' gestures, all of which are low speed gestures. RC3 had high loadings on the 'Bored', 'Cautious', 'Dance natural motion', and 'Look Around natural motion' gestures. Only two gestures had high factor loadings on RC4, which makes it difficult to identify.

The principal components analysis for elicited arousal response resulted in two rotated components which accounted for 30% and 25% of the total variance respectively. All of the gestures that load highly on the first component could be qualitatively described as slow gestures, and all of the gestures that load highly on the second component can be described as fast gestures. The three 'Point' gestures at 50%, 75%, and 100% speed had high loadings on the second component. This leads us to believe that the first factor of arousal is 'melancholy' and the second 'aggressive' or 'high velocity' motions.

The principal components analysis for the perceived valence responses resulted in five rotated components which accounted for 15%, 13%, 13%, 12%, and 12% respectively. Unfortunately, the results are perhaps to muddled to interpret, with only half of the gestures having a high factor loading.

The principal components analysis for the perceived arousal responses resulted in four rotated components which accounted for 17%, 15%, 12%, and 11% of the total variation respectively. The first rotated component, RC1, shows high factor loading for the three high speed 'Point' gestures, which strongly suggests that RC1 is related to high speed. RC2 shows a high loading on 'Bored', 'Desolation', 'Taunt', and 'Wave'. These gestures are all low speed gestures which suggests RC2 is related to low speed. RC3 shows high factor loadings for 'Cautious' and 'natural motion Look Around' both of which have sudden jerky movements, suggesting RC3 is related to smoothness. RC4 shows high loading for 'Curiosity natural motion', 'Dance natural motion', and 'Present'. These three gestures all have a large scale or TSR, which suggests RC4 is related to gesture size.

Responses by Demographics

If there are any differences in responses due to subjects’ demographic groups, this could provide valuable insight into how a socially embedded robot may want to change its gestures when interacting with person(s) from those demographic groups. As an example, if Male subjects have higher elicited excitement responses due to certain kinematic features such as speed or smoothness of motion and the desired goal of an instance of human-robot interaction is to calm the subject, the designer of that robot may want to reduce the speed of the gestures the robot uses.

To determine whether or not there was a significant difference in the responses given we first determined whether the variance was similar for responses by females and males. A T-test was then conducted for each gesture to determine whether or not the samples came from the same underlying population, that is whether or not males and females had different responses. As can be seen in figure 14, only certain gestures are likely drawn from different populations.

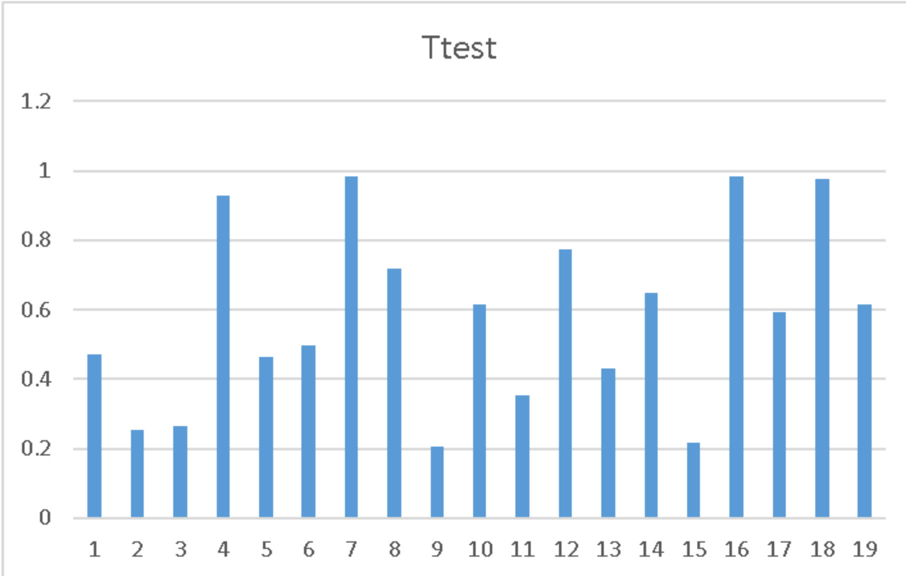


Figure 14 - T test results; probability that male/female responses are drawn from the same mean, per gesture

Across all 4 questions and 19 gestures there were only 14 instances where the probabilities of responses being drawn from the same population was below .20, as can be seen in table 2. In fact, the ‘Elicited Arousal’ responses were likely drawn from the same population indicating that males and females both reported similarly when asked how excited the gesture made them feel.

Table 4 Results of T-Test

	Gesture																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Elicited Valence	0.22	0.73	0.57	0.16	0.20	0.16	0.08	0.36	0.32	0.39	0.02	0.30	0.40	0.38	0.60	0.29	0.52	0.12	0.07
Elicted Arousal	0.47	0.25	0.27	0.93	0.46	0.50	0.98	0.72	0.21	0.62	0.35	0.77	0.43	0.65	0.22	0.98	0.59	0.98	0.61
Percieved Valenc	0.35	0.27	0.90	0.75	1.00	0.45	0.86	0.86	0.94	0.19	0.05	0.55	0.92	0.60	0.01	0.22	0.17	0.92	0.65
Percieved Arousa	0.13	0.26	0.14	0.91	0.83	0.47	0.56	0.72	0.72	0.31	0.24	0.90	0.15	0.76	0.46	1.00	0.49	0.03	0.28

Gesture Lexicon

As mentioned previously, a key objective of this project was the development and publishing of a lexicon of movements effective in communicating ideas and emotions to humans. We hope as this library continues to develop, it will prove an invaluable tool in studying and improving HRI.

In classifying gestures, we analyzed their kinematic properties and their common intents, categorizing them in any of the four movement types listed previously. Following the Task-Space Region template, we also found it useful to track movement of the end-effector exclusively. We theorized that this would correlate to interactors' comfort zone and overall confidence in the experiment.

TSRs

In calculating the TSRs using the final RAPID code export data, there were several steps to generating clean and precise data. Our code takes a list of jtags containing angles for each of the six joints on the arm. The data is read as [deltaT, angleA, angleB, angleC, angleD, angleE, angleF], with each time-stamped frame separated by commas. We wrote a simple python script to read in each .txt file and output the calculated values for a cartesian representation of the end-effector's location in time.

The python function took angleB and angle C (keeping in mind that the rapid code angles for joint C is dependent on those of joint B, giving them independent reference frames) and output the j and k components of a cartesian vector. This system assumes that the robot zero position, that is, [deltaT, 0, 0, 0, 0, 0, 0], is as shown in figure 15.

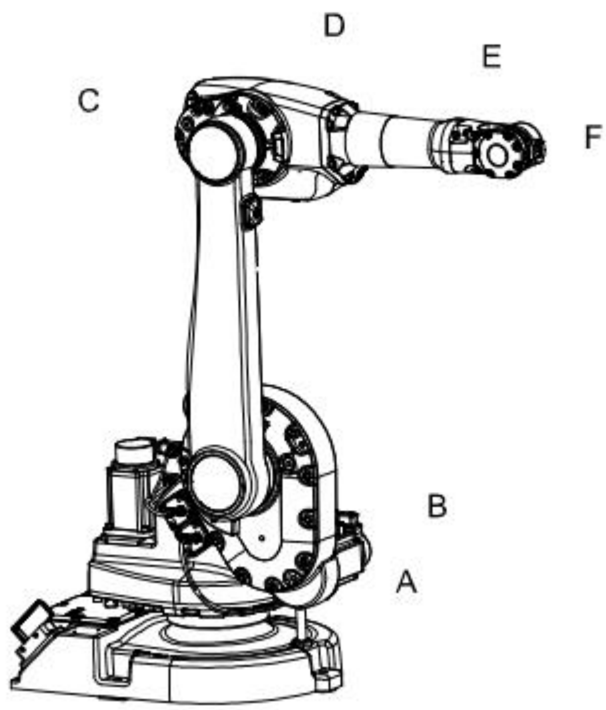


Figure 15 - Joint identification and robot zero-position

As shown in figure 15, the robot zero-position keeps link BC at a 90 degree angle from a global y-axis, while link CD is held at a 0-degree angle from the global y-axis. This showed the angles to have the following simple relationship.

$$angC = angB + 90$$

$$angB = globB + 90$$

Where globB is the angle that link BC makes with the horizontal y –axis. Therefore, the global C angle was defined as:

$$globC = angC + globB - 90$$

Once the angles were transformed into a global reference frame, it was simple to use basic trigonometry to calculate the Cartesian position using the dimensions of the robot with vector addition.

Once the script printed values, we output them to a functional excel spreadsheet to confirm the math, and took the max i, j, and k values for the x, y, and z components of the TSR. One of the interesting quantities we thought would be interesting was the normalized distance from our local origin, that is, the base of the robot's joint A. The figures below compare two versions of the curiosity gesture, one motion-captured using our HTC Vive, and one traced using our Blender plugin.

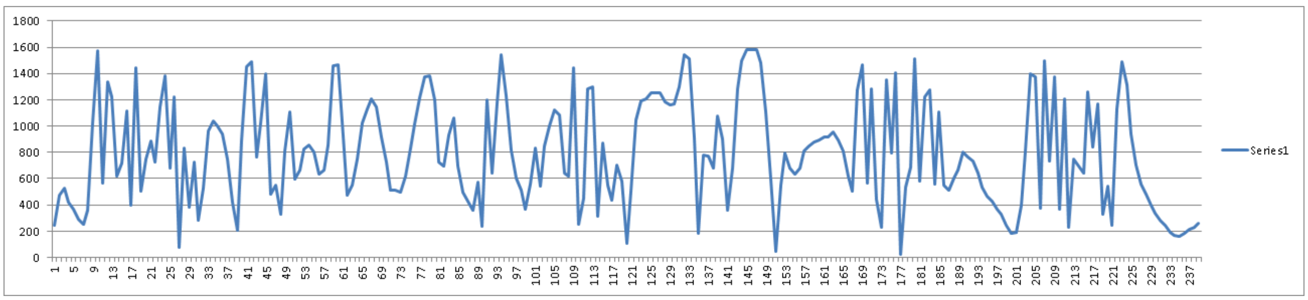


Figure 16: Mo-capped Curiosity Gesture

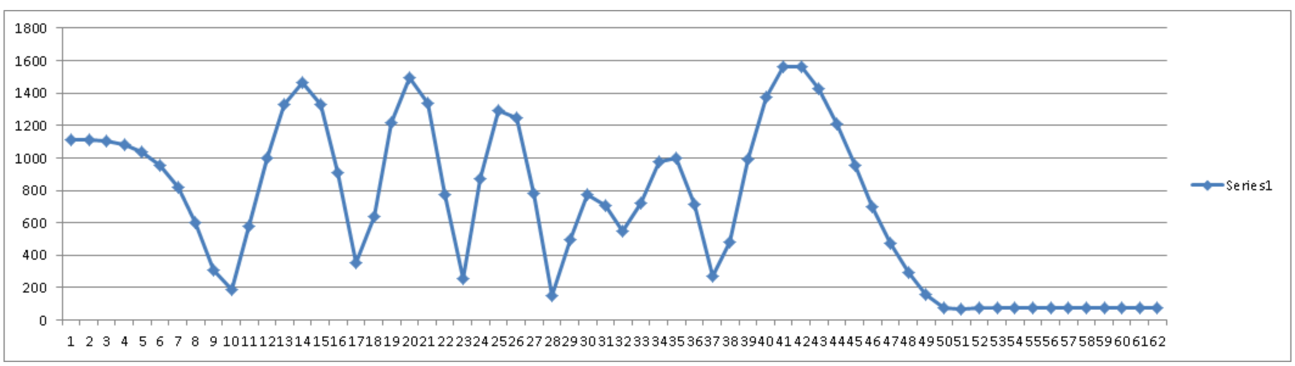


Figure 17: Blender-traced Curiosity Gesture

Obviously, the motion-captured data shows much less smooth and consistent data, picking up twitches in movement exhibited by a shaky human actor. The displayed graphs essentially show the change in size of the effective space in which the robot is active at that point in time. This quantity over time demonstrates the wide range of the end-effector's movement, but it does not provide a precise boundary value for the operating space of each gesture. With the TSR's we are able to demonstrate the quantity of motion that the end-effector exhibits in each case. The TSR for each gesture is given as a min/max range for the X, Y, and Z global axes, and can be seen as part of the lexicon master document, attached to our report.

Gesture Velocity Analysis:

Along with Calculating the TSR's for each gesture the team also calculated the velocities over time for each one. This was done by taking the joint positions array used to animation and turning them into an excel spreadsheet, after that MATLAB was used to graph the angles and angular velocity of each joint. The idea being that gestures with high velocities will result in the gestures bring more uncomfortable to watch.

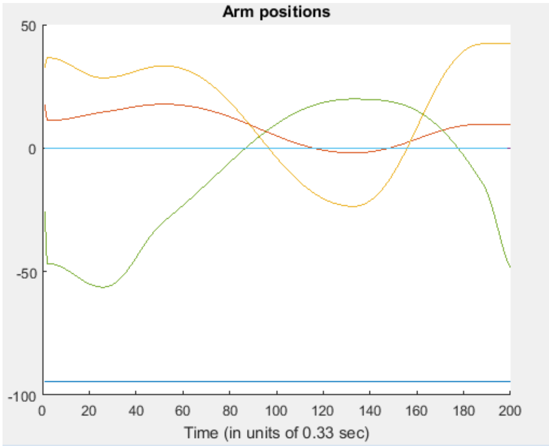


Figure 18: Joint Angles

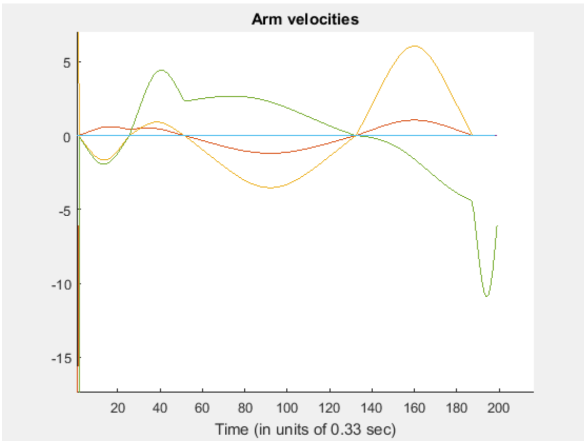


Figure 19 Joint Velocities

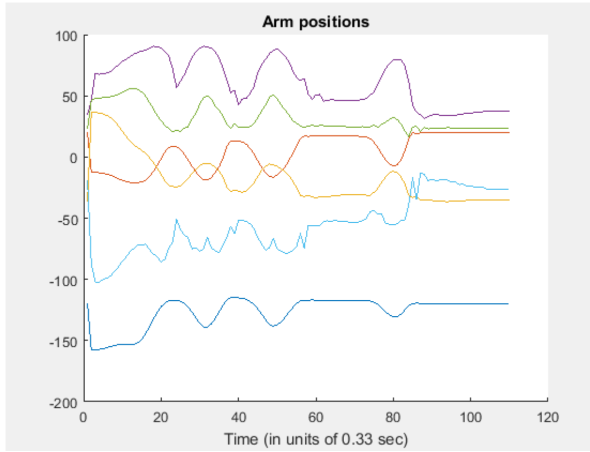


Figure 20

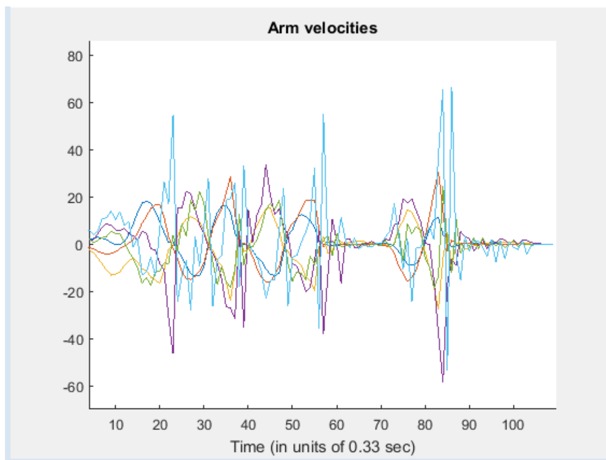


Figure 21

In the figures above one can see the difference between a low velocity gesture (boredom) and a high velocity gesture (Indication). With each different colored line representing a different joint of the robot. After looking at the experiment results the participants rated the high velocity gestures as being far more displaceable compared to the low velocities gestures. The high velocities gestures also appeared to be more energetic with the participants rating them high on the arousal scale. This these results seemed to conform to the original postulate that that faster gestures are more uncomfortable to watch. It is not only high velocity gestures that are seen as uncomfortable but it is also gestures that have rapid changes in velocity and direction, jerky movements. These types of movements could be observed in gestures generated with motion capture data, contrasting the manually animated gestures that have been synthetically smoothed out This however is not taking into account the noises the robot makes at high speeds which might be disturbing to the participants.

Conclusions & Recommendations

We used the term Creative Robotics to generally refer to robotic systems in artistic senses or environments. Indeed, many interesting projects have explored the use of robots in performances such as dances or symphonies. In our case, we decided to focus on the fundamentals of art; that is, the potential to elicit emotion through expression.

In many artistic expressions, there are some form of gesture or meaningful motion. When singing, dancing, or acting, there are implicit and explicit body languages. In presentations or speeches, performers commonly emphasize points with hand motions. As stated previously, we designed this experiment to analyze the kinematic quantities of these motions in gesture analysis.

Our findings overall were decidedly less significant than we had hoped. However, with this experiment we hoped to define methodologies to more effectively experiment in creative robotics. That said, a number of interesting correlations were shown between quantities like TSR and comfort zone, as well as those in the preliminary survey results.

With smaller sample sizes and limited time to execute our experiment, our conclusions are primarily focused on improving the experiment methodology and data collection methods.

After the preliminary survey and the final experiment, we received a lot of comments and feedback about our execution of certain things, including question phrasing and lab setup. One student asked if we had considered where observers were standing, indicating that each respondent to the questionnaire had a different view of the displayed gesture. This issue was something we had considered but unfortunately did not have time to integrate into the experiment. Another student, upon receiving our debriefing document, voiced their opinion on our hypotheses, stating their unease with higher-velocity and larger-TSR gestures, while identifying smaller, slower gestures as relaxed and even “cute”.

Future Works:

A future experiment could improve upon this experiment in at least three ways. As mentioned before with fifteen subjects standing around the robot not all subjects were facing the front of the robot which was the area the gestures were designed to be viewed from. Another way in which future works could improve upon this experiment is by having more subjects participate in the experiment. The third way in which a future experiment could improve upon this experiment is by have a single method of gesture creation.

The space in which the robot is situated allowed for a maximum of fifteen participants to view the robot at a time. However, this number of participants at once prevented all participants from viewing from the front of the robot. Since the gestures were designed to be viewed from the front some elements of the gesture could have been missed when viewing from the side. Additionally, since the viewpoint of each participant was not tracked, this introduces an unknown variable that could change the validity of results. Ideally a smaller number of people would be brought into the experiment at a time to reduce the problem of location of subjects having different views. This would be possible since we had estimated one-hour time slots and had finished each trial within thirty minutes.

The minimum number of participants for the principle components analysis to have significance was fifty. However, we only had thirty participants. This limited the capability to determine underlying factors. The sign up for the experiment was not announced until within a week of the experiment time. Additionally, the experiment was at 10 am on a Saturday. On the campus of the university early morning on Saturday is not normally a highly active time. Both of these problems are due to late scheduling of the experiment. By the time we went to schedule the experiment most of the ideal time slots were already taken. As noted above each individual trial could be shorter which might encourage more participants as well as allow for more flexibility in time slot scheduling. Instead of scheduling a three-hour block, multiple shorter blocks could be scheduled. This would increase the chance that potential participates would be available to participate.

As mentioned above some gestures were animated by hand and others used motion capture data from an HTC Vive. The natural motion gestures tended to have more key frames in development. Additionally, the natural gestures contained the actor’s hand shaking on a minor level. This made the gestures contain a bit of shaking that was not present in the hand animated gestures. At the same time, they more closely simulated how the gesture would be performed by a human. A future work could use the already established code for turning data from an HTC Vive into blender animations as well as the code from the previous group that allowed blender animations to be turned into rapid code to animate gestures more quickly.

We designing the gestures to use for the performance, most gestures were made to mimic a particular gesture of a human. As a result, many of the gestures had underlying context to the particulars of the motion. This resulted in the suspected factors not having proper variations that held the other factors constant. The point gesture was run at several speeds to test if a variation of speed keeping other factors constant would cause a change in responses. One potential way to better represent the suspected factors of trajectory, speed, and acceleration would be to have a set number base gestures that have different trajectories and run each with the same number of variations in speed and acceleration. For instance, if five gestures were to be created then each one would be run with five speeds and five accelerations leading to 125 different animations to run. This would better represent the effective space of the suspected factors.

References

Berenson, D., Srinivasa, S., & Kuffner, J. (2011). Task Space Regions: A framework for pose-constrained manipulation planning. *The International Journal of Robotics Research*, *30*(12), 1435–1460. https://doi.org/10.1177/0278364910396389

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*(1–2), 119–155. https://doi.org/10.1016/S1071-5819(03)00018-1

Bremner, P., & Leonards, U. (2016). Iconic Gestures for Robot Avatars, Recognition and Integration with Speech. *Human-Media Interaction*, 183. https://doi.org/10.3389/fpsyg.2016.00183

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, *59*(1–2), 213–225. https://doi.org/10.1016/S1071-5819(03)00050-8

Cassell, J., Vilhjálmsson, H. H., & Bickmore, T. (2004). BEAT: The Behavior Expression Animation Toolkit. In H. Prendinger & M. Ishizuka (Eds.), *Life-Like Characters* (pp. 163–185). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-662-08373-4_8

Castellano, G., Villalba, S. D., & Camurri, A. (2007). Recognising Human Emotions from Body Movement and Gesture Dynamics. In A. C. R. Paiva, R. Prada, & R. W. Picard (Eds.), *Affective Computing and Intelligent Interaction* (pp. 71–82). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-74889-2_7

Daly, E. M., Lancee, W. J., & Polivy, J. (1983). A conical model for the taxonomy of emotional experience. *Journal of Personality and Social Psychology*, *45*(2), 443–457. https://doi.org/10.1037/0022-3514.45.2.443

Dautenhahn, K., Ogden, B., & Quick, T. (2002). From embodied to socially embedded agents – Implications for interaction-aware robots. *Cognitive Systems Research*, *3*(3), 397–428. https://doi.org/10.1016/S1389-0417(02)00050-5

Dragan, A. D., Bauman, S., Forlizzi, J., & Srinivasa, S. S. (2015). Effects of Robot Motion on Human-Robot Collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 51–58). New York, NY, USA: ACM. https://doi.org/10.1145/2696454.2696473

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*(3–4), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

Gleeson, B., MacLean, K., Haddadi, A., Croft, E., & Alcazar, J. (2013). Gestures for Industry: Intuitive Human-robot Communication from Human Observation. In *Proceedings of the 8th ACM/IEEE*

International Conference on Human-robot Interaction (pp. 349–356). Piscataway, NJ, USA: IEEE Press. Retrieved from http://dl.acm.org/citation.cfm?id=2447556.2447679

History | iRobot. (n.d.). Retrieved October 4, 2016, from http://www.irobot.com/About-iRobot/Company-Information/History.aspx

Holladay, R., & Srinivasa, S. S. (2016). RoGuE. Robotics Institute, Carnegie Mellon University. Retrieved from https://www.ri.cmu.edu/pub_files/2016/3/rogue.pdf

Huang, C.-M., & Mutlu, B. (n.d.). Modeling and Evaluating Narrative Gestures for Humanlike Robots. Department of Computer Sciences, University of Wisconsin.

Knight, W. (n.d.). This Robot Could Transform Manufacturing. Retrieved May 7, 2017, from https://www.technologyreview.com/s/429248/this-robot-could-transform-manufacturing/

Kopp, S., & Wachsmuth, I. (2000). A Knowledge-based Approach for Lifelike Gesture Animation. In *Proceedings of the 14th European Conference on Artificial Intelligence* (pp. 663–667). Amsterdam, The Netherlands, The Netherlands: IOS Press. Retrieved from http://dl.acm.org/citation.cfm?id=3006433.3006574

Kosuge, K., & Hirata, Y. (2004). Human-Robot Interaction. In *IEEE International Conference on Robotics and Biomimetics, 2004. ROBIO 2004* (pp. 8–11). https://doi.org/10.1109/ROBIO.2004.1521743

Laguillaumie, P., Laribi, M. A., Seguin, P., Vulliez, P., Decatoire, A., & Zeghloul, S. (2016). From Human Motion Capture to Industrial Robot Imitation. In S. Zeghloul, M. A. Laribi, & J.-P. Gazeau (Eds.), *Robotics and Mechatronics* (pp. 301–312). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-22368-1_30

Lalwani, M. (n.d.). Lady Gaga’s robotic keyboard had some help from NASA. Retrieved May 7, 2017, from https://www.engadget.com/2016/02/20/lady-gaga-robotic-keyboard-grammys-nasa/

Li, J., & Chignell, M. (2010). Communication of Emotion in Social Robots through Simple Head and Arm Movements. *International Journal of Social Robotics*, *3*(2), 125–142. https://doi.org/10.1007/s12369-010-0071-x

Lu, D. V., & Smart, W. D. (2011). Human-robot interactions as theatre. In *2011 RO-MAN* (pp. 473–478). https://doi.org/10.1109/ROMAN.2011.6005241

Mehrabian, A. (n.d.). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. *Current Psychology*, *14*(4), 261–292. https://doi.org/10.1007/BF02686918

Misselhorn, C. (2009). Empathy with Inanimate Objects and the Uncanny Valley. *Minds and Machines*, *19*(3), 345. https://doi.org/10.1007/s11023-009-9158-2

Nam, H. Y., & Choi, C. (2012). Artistic Robot Please Smile. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 967–970). New York, NY, USA: ACM. https://doi.org/10.1145/2212776.2212362

Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. In *2011 IEEE International Conference on Robotics and Automation* (pp. 3400–3407). https://doi.org/10.1109/ICRA.2011.5979561

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (1989). Measures of emotion. In R. Plutchik & H. Kellerman (Eds.), *The measurement of emotions* (pp. 83–111). San Diego, CA, US: Academic Press.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(3), 273–294. https://doi.org/10.1016/0092-6566(77)90037-X

Saerbeck, M., & Bartneck, C. (2010). Perception of Affect Elicited by Robot Motion. In *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction* (pp. 53–60). Piscataway, NJ, USA: IEEE Press. Retrieved from http://dl.acm.org/citation.cfm?id=1734454.1734473

Salem, M., Kopp, S., Wachsmuth, I., Rohlfing, K., & Joublin, F. (2012). Generation and Evaluation of Communicative Robot Gesture. *International Journal of Social Robotics*, *4*(2), 201–217. https://doi.org/10.1007/s12369-011-0124-9

Sawada, M., Suda, K., & Ishii, M. (2003). Expression of Emotions in Dance: Relation between Arm Movement Characteristics and Emotion. *Perceptual and Motor Skills*, *97*(3), 697–708. https://doi.org/10.2466/pms.2003.97.3.697

Shikanai, N., Sawada, M., & Ishii, M. (2013). Development of the Movements Impressions Emotions Model: Evaluation of Movements and Impressions Related to the Perception of Emotions in Dance. *Journal of Nonverbal Behavior*, *37*(2), 107–121. https://doi.org/10.1007/s10919-013-0148-y

Sial, S. B., Sial, M. B., Ayaz, Y., Shah, S. I. A., & Zivanovic, A. (2016). Interaction of robot with humans by communicating simulated emotional states through expressive movements. *Intelligent Service Robotics*, *9*(3), 231–255. https://doi.org/10.1007/s11370-016-0199-0

Stanton, C., & Stevens, C. J. (2014). Robot Pressure: The Impact of Robot Eye Gaze and Lifelike Bodily Movements upon Decision-Making and Trust. In M. Beetz, B. Johnston, & M.-A. Williams

(Eds.), *Social Robotics* (pp. 330–339). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-11973-1_34

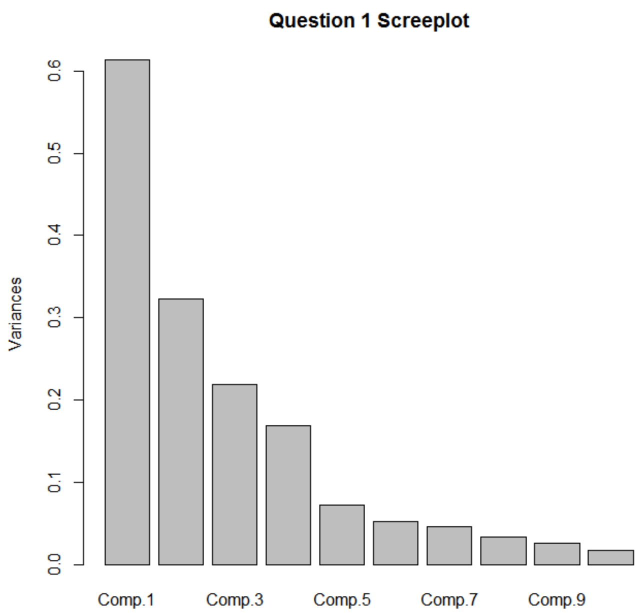
Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006). Common Metrics for Human-robot Interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction* (pp. 33–40). New York, NY, USA: ACM. https://doi.org/10.1145/1121241.1121249

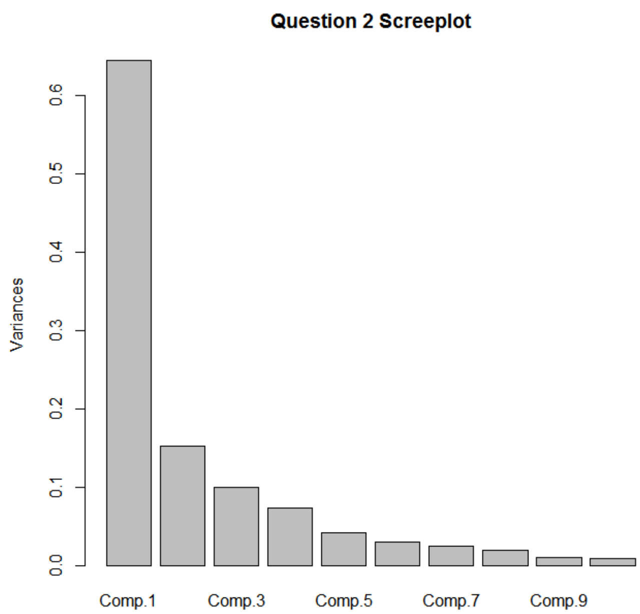
Stevens, C. J., Schubert, E., Morris, R. H., Frear, M., Chen, J., Healey, S., … Hansen, S. (2009). Cognition and the temporal arts: Investigating audience response to dance using PDAs that record continuous data during live performance. *International Journal of Human-Computer Studies*, *67*(9), 800–813. https://doi.org/10.1016/j.ijhcs.2009.06.001

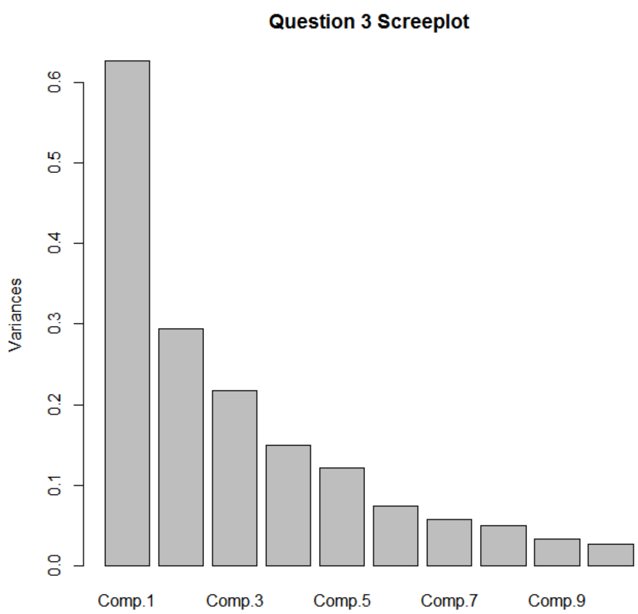
Tonietti, G., Schiavi, R., & Bicchi, A. (2005). Design and Control of a Variable Stiffness Actuator for Safe and Fast Physical Human/Robot Interaction. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation* (pp. 526–531). https://doi.org/10.1109/ROBOT.2005.1570172

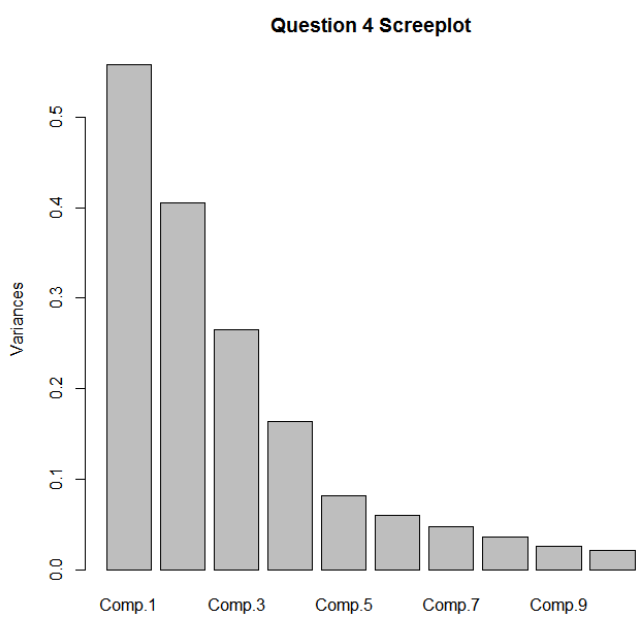
Wheelock, C. (n.d.). Healthcare Robot Shipments to Surpass 10,000 Units Annually by 2021 | Tractica. Retrieved March 21, 2017, from https://www.tractica.com/newsroom/press-releases/healthcare-robot-shipments-to-surpass-10000-units-annually-by-2021/

Appendix A – PCA Results









Question 1 PCA Results:

principal(r=cor(data),nfactors=4,rotate="varimax",n.obs=27)

Principal Components Analysis

Call: principal(r = cor(data), nfactors = 4, rotate = "varimax", n.obs = 27)

Standardized loadings (pattern matrix) based upon correlation matrix

	RC1	RC2	RC3	RC4	h2	u2	com
Gesture.1	-0.13	0.07	0.69	-0.26	0.56	0.44	1.4
Gesture.2	0.01	0.05	0.65	0.39	0.58	0.42	1.7
Gesture.3	-0.20	0.33	0.28	0.25	0.29	0.71	3.5
Gesture.4	0.48	0.08	0.39	-0.20	0.43	0.57	2.4
Gesture.5	0.28	-0.34	0.75	0.00	0.76	0.24	1.7
Gesture.6	0.08	0.17	0.10	0.65	0.47	0.53	1.2
Gesture.7	-0.12	-0.03	0.14	-0.52	0.31	0.69	1.3
Gesture.8	0.25	0.26	0.68	0.20	0.63	0.37	1.8
Gesture.9	0.03	0.86	-0.09	0.08	0.75	0.25	1.0
Gesture.10	0.27	0.78	-0.16	-0.04	0.71	0.29	1.3
Gesture.11	0.67	0.03	0.38	-0.05	0.60	0.40	1.6
Gesture.12	-0.05	0.87	0.08	0.36	0.89	0.11	1.4
Gesture.13	-0.40	0.10	0.28	0.64	0.65	0.35	2.2
Gesture.14	-0.07	0.63	0.29	-0.01	0.49	0.51	1.4
Gesture.15	-0.08	0.48	0.47	-0.46	0.67	0.33	3.0
Gesture.16	-0.60	0.26	0.07	-0.30	0.52	0.48	1.9
Gesture.17	0.68	0.08	-0.18	-0.04	0.51	0.49	1.2
Gesture.18	0.91	0.17	0.01	-0.04	0.86	0.14	1.1
Gesture.19	0.80	-0.12	0.23	0.23	0.76	0.24	1.4

	RC1	RC2	RC3	RC4
SS loadings	3.44	3.19	2.84	1.95
Proportion Var	0.18	0.17	0.15	0.10
Cumulative Var	0.18	0.35	0.50	0.60
Proportion Explained	0.30	0.28	0.25	0.17
Cumulative Proportion	0.30	0.58	0.83	1.00

Mean item complexity = 1.7

Test of the hypothesis that 4 components are sufficient.

The root mean square of the residuals (RMSR) is 0.09

with the empirical chi square 82.68 with prob < 0.91

Fit based upon off diagonal values = 0.86

Question 2 PCA Results

```
> principal(r=cor(data2),nfactors=2,rotate="varimax",n.obs=26)
Principal Components Analysis
Call: principal(r = cor(data2), nfactors = 2, rotate = "varimax", n.obs = 26)
Standardized loadings (pattern matrix) based upon correlation matrix
            RC1   RC2   h2   u2 com
Gesture.1   0.64  0.30 0.50 0.50 1.4
Gesture.2   0.66  0.15 0.46 0.54 1.1
Gesture.3   0.52  0.39 0.42 0.58 1.9
Gesture.4   0.15  0.75 0.59 0.41 1.1
Gesture.5   0.06  0.71 0.51 0.49 1.0
Gesture.6   0.84  0.03 0.71 0.29 1.0
Gesture.7   0.56  0.21 0.36 0.64 1.3
Gesture.8   0.61  0.26 0.45 0.55 1.3
Gesture.9   0.62  0.00 0.39 0.61 1.0
Gesture.10  0.56  0.52 0.57 0.43 2.0
Gesture.11  0.16  0.63 0.42 0.58 1.1
Gesture.12  0.41  0.36 0.30 0.70 2.0
Gesture.13  0.67  0.00 0.45 0.55 1.0
Gesture.14  0.60  0.30 0.45 0.55 1.5
Gesture.15  0.85 -0.05 0.73 0.27 1.0
Gesture.16  0.83 -0.14 0.71 0.29 1.1
Gesture.17 -0.05  0.88 0.77 0.23 1.0
Gesture.18  0.10  0.86 0.75 0.25 1.0
Gesture.19  0.09  0.91 0.83 0.17 1.0

                       RC1  RC2
SS loadings           5.69 4.69
Proportion Var        0.30 0.25
Cumulative Var        0.30 0.55
Proportion Explained  0.55 0.45
Cumulative Proportion 0.55 1.00
```

Mean item complexity = 1.3

Test of the hypothesis that 2 components are sufficient.

The root mean square of the residuals (RMSR) is 0.11

with the empirical chi square 103.16 with prob < 0.98

Fit based upon off diagonal values = 0.92

Question 3 PCA Results

> principal(r=cor(data3),nfactors=5,rotate="varimax",n.obs=26)

Principal Components Analysis

Call: principal(r = cor(data3), nfactors = 5, rotate = "varimax", n.obs = 26)

Standardized loadings (pattern matrix) based upon correlation matrix

	RC4	RC5	RC2	RC3	RC1	h2	u2	com
Gesture.1	0.12	0.78	-0.06	0.15	-0.15	0.68	0.32	1.2
Gesture.2	0.68	0.02	0.05	0.32	0.16	0.59	0.41	1.6
Gesture.3	0.03	0.45	0.62	0.24	0.16	0.68	0.32	2.3
Gesture.4	0.88	-0.03	-0.07	-0.12	-0.02	0.79	0.21	1.1
Gesture.5	-0.10	0.06	-0.74	0.33	0.18	0.70	0.30	1.6
Gesture.6	-0.08	-0.08	0.72	0.01	0.24	0.59	0.41	1.3
Gesture.7	0.01	-0.83	0.09	0.08	-0.29	0.78	0.22	1.3
Gesture.8	0.84	0.11	-0.02	0.17	-0.05	0.75	0.25	1.1
Gesture.9	0.20	0.09	0.26	-0.01	0.77	0.71	0.29	1.4
Gesture.10	-0.21	0.16	0.05	0.17	0.73	0.63	0.37	1.4
Gesture.11	0.13	-0.04	-0.16	0.69	0.27	0.59	0.41	1.5
Gesture.12	0.54	0.03	0.33	-0.20	0.43	0.63	0.37	3.0
Gesture.13	0.12	0.50	0.30	-0.35	0.27	0.55	0.45	3.3
Gesture.14	0.18	0.35	0.40	0.59	0.14	0.68	0.32	2.8
Gesture.15	0.31	-0.01	-0.18	-0.46	0.13	0.36	0.64	2.3
Gesture.16	-0.17	0.15	0.20	-0.74	0.07	0.65	0.35	1.4
Gesture.17	0.27	-0.25	-0.40	-0.08	0.54	0.59	0.41	3.0
Gesture.18	0.17	-0.57	-0.44	0.31	0.24	0.70	0.30	3.1
Gesture.19	0.37	-0.40	-0.18	0.32	0.44	0.62	0.38	4.2

	RC4	RC5	RC2	RC3	RC1
SS loadings	2.77	2.50	2.40	2.32	2.29
Proportion Var	0.15	0.13	0.13	0.12	0.12

Cumulative Var 0.15 0.28 0.40 0.53 0.65

Proportion Explained 0.23 0.20 0.20 0.19 0.19

Cumulative Proportion 0.23 0.43 0.63 0.81 1.00

Mean item complexity = 2

Test of the hypothesis that 5 components are sufficient.

The root mean square of the residuals (RMSR) is 0.09

with the empirical chi square 74.7 with prob < 0.8

Fit based upon off diagonal values = 0.86>

Question 4 Results

> principal(r=cor(data4),nfactors=4,rotate="varimax",n.obs=26)

Principal Components Analysis

Call: principal(r = cor(data4), nfactors = 4, rotate = "varimax", n.obs = 26)

Standardized loadings (pattern matrix) based upon correlation matrix

	RC1	RC2	RC3	RC4	h2	u2	com
Gesture.1	0.02	0.63	0.39	0.01	0.55	0.45	1.7
Gesture.2	-0.35	-0.21	0.77	-0.07	0.77	0.23	1.6
Gesture.3	0.26	0.49	0.13	0.03	0.33	0.67	1.7
Gesture.4	0.08	0.03	0.26	0.68	0.54	0.46	1.3
Gesture.5	0.21	-0.13	-0.02	0.72	0.58	0.42	1.2
Gesture.6	-0.05	0.63	-0.34	-0.25	0.58	0.42	1.9
Gesture.7	-0.45	-0.24	-0.28	0.51	0.60	0.40	3.0
Gesture.8	-0.22	-0.06	0.79	0.13	0.70	0.30	1.2
Gesture.9	-0.67	0.10	0.14	0.17	0.50	0.50	1.3
Gesture.10	0.46	0.37	0.37	-0.03	0.49	0.51	2.9
Gesture.11	0.07	0.23	-0.25	0.65	0.54	0.46	1.6
Gesture.12	0.18	0.06	0.44	-0.32	0.33	0.67	2.2
Gesture.13	0.25	0.29	-0.18	-0.19	0.22	0.78	3.4
Gesture.14	-0.10	0.67	0.02	-0.12	0.48	0.52	1.1
Gesture.15	-0.11	0.71	-0.32	0.10	0.62	0.38	1.5
Gesture.16	-0.20	0.59	-0.06	0.17	0.42	0.58	1.4
Gesture.17	0.82	-0.05	-0.24	0.14	0.76	0.24	1.2
Gesture.18	0.65	-0.24	-0.11	0.28	0.58	0.42	1.7
Gesture.19	0.85	0.01	0.06	0.16	0.76	0.24	1.1

	RC1	RC2	RC3	RC4
SS loadings	3.14	2.82	2.28	2.10
Proportion Var	0.17	0.15	0.12	0.11
Cumulative Var	0.17	0.31	0.43	0.54
Proportion Explained	0.30	0.27	0.22	0.20
Cumulative Proportion	0.30	0.58	0.80	1.00

Mean item complexity = 1.7
Test of the hypothesis that 4 components are sufficient.

The root mean square of the residuals (RMSR) is 0.11
with the empirical chi square 101.55 with prob < 0.47

Fit based upon off diagonal values = 0.77

Question 1 Loadings

		RC1	RC2	RC3	RC4
Gesture.1	Bored	-0.13	0.07	0.69	-0.26
Gesture.2	Cautious	0.01	0.05	0.65	0.39
Gesture.3	Curiosity	-0.2	0.33	0.28	0.25
Gesture.4	Curiousity_NM	0.48	0.08	0.39	-0.2
Gesture.5	Dance_NM	0.28	-0.34	0.75	0
Gesture.6	Desolation	0.08	0.17	0.1	0.65
Gesture.7	Excited	-0.12	-0.03	0.14	-0.52
Gesture.8	LookAround_NM	0.25	0.26	0.68	0.2
Gesture.9	No	0.03	0.86	-0.09	0.08
Gesture.10	Point_NM_25	0.27	0.78	-0.16	-0.04
Gesture.11	Present	0.67	0.03	0.38	-0.05
Gesture.12	Present 2_NM	-0.05	0.87	0.08	0.36
Gesture.13	Stop	-0.4	0.1	0.28	0.64
Gesture.14	Taunt	-0.07	0.63	0.29	-0.01
Gesture.15	Wave	-0.08	0.48	0.47	-0.46
Gesture.16	Yes	-0.6	0.26	0.07	-0.3
Gesture.17	Point_NM_50	0.68	0.08	-0.18	-0.04
Gesture.18	Point_NM_75	0.91	0.17	0.01	-0.04
Gesture 19	Point_NM_100	0.8	-0.12	0.23	0.23

Question 2 Loadings

		RC1	RC2
Gesture.1	Bored	0.64	0.3
Gesture.2	Cautious	0.66	0.15
Gesture.3	Curiosity	0.52	0.39
Gesture.4	Curiousity_NM	0.15	0.75
Gesture.5	Dance_NM	0.06	0.71
Gesture.6	Desolation	0.84	0.03
Gesture.7	Excited	0.56	0.21
Gesture.8	LookAround_NM	0.61	0.26
Gesture.9	No	0.62	0
Gesture.10	Point_NM_25	0.56	0.52
Gesture.11	Present	0.16	0.63
Gesture.12	Present 2_NM	0.41	0.36
Gesture.13	Stop	0.67	0
Gesture.14	Taunt	0.6	0.3
Gesture.15	Wave	0.85	-0.05
Gesture.16	Yes	0.83	-0.14
Gesture.17	Point_NM_50	-0.05	0.88
Gesture.18	Point_NM_75	0.1	0.86
Gesture.19	Point_NM_100	0.09	0.91

Question 3 Loadings

		RC1	RC2	RC3	RC4	RC5
Gesture.1	Bored	0.12	0.78	-0.06	0.15	-0.15
Gesture.2	Cautious	0.68	0.02	0.05	0.32	0.16
Gesture.3	Curiosity	0.03	0.45	0.62	0.24	0.16
Gesture.4	Curiousity_NM	0.88	-0.03	-0.07	-0.12	-0.02
Gesture.5	Dance_NM	-0.1	0.06	-0.74	0.33	0.18
Gesture.6	Desolation	-0.08	-0.08	0.72	0.01	0.24
Gesture.7	Excited	0.01	-0.83	0.09	0.08	-0.29
Gesture.8	LookAround_NM	0.84	0.11	-0.02	0.17	-0.05
Gesture.9	No	0.2	0.09	0.26	-0.01	0.77
Gesture.10	Point_NM_25	-0.21	0.16	0.05	0.17	0.73
Gesture.11	Present	0.13	-0.04	-0.16	0.69	0.27
Gesture.12	Present 2_NM	0.54	0.03	0.33	-0.2	0.43
Gesture.13	Stop	0.12	0.5	0.3	-0.35	0.27
Gesture.14	Taunt	0.18	0.35	0.4	0.59	0.14
Gesture.15	Wave	0.31	-0.01	-0.18	-0.46	0.13
Gesture.16	Yes	-0.17	0.15	0.2	-0.74	0.07
Gesture.17	Point_NM_50	0.27	-0.25	-0.4	-0.08	0.54
Gesture.18	Point_NM_75	0.17	-0.57	-0.44	0.31	0.24
Gesture.19	Point_NM_100	0.37	-0.4	-0.18	0.32	0.44

Question 4 Loadings

		RC1	RC2	RC3	RC4
Gesture.1	Bored	0.02	0.63	0.39	0.01
Gesture.2	Cautious	-0.35	-0.21	0.77	-0.07
Gesture.3	Curiosity	0.26	0.49	0.13	0.03
Gesture.4	Curiousity_NM	0.08	0.03	0.26	0.68
Gesture.5	Dance_NM	0.21	-0.13	-0.02	0.72
Gesture.6	Desolation	-0.05	0.63	-0.34	-0.25
Gesture.7	Excited	-0.45	-0.24	-0.28	0.51
Gesture.8	LookAround_NM	-0.22	-0.06	0.79	0.13
Gesture.9	No	-0.67	0.1	0.14	0.17
Gesture.10	Point_NM_25	0.46	0.37	0.37	-0.03
Gesture.11	Present	0.07	0.23	-0.25	0.65
Gesture.12	Present 2_NM	0.18	0.06	0.44	-0.32
Gesture.13	Stop	0.25	0.29	-0.18	-0.19
Gesture.14	Taunt	-0.1	0.67	0.02	-0.12
Gesture.15	Wave	-0.11	0.71	-0.32	0.1
Gesture.16	Yes	-0.2	0.59	-0.06	0.17
Gesture.17	Point_NM_50	0.82	-0.05	-0.24	0.14
Gesture.18	Point_NM_75	0.65	-0.24	-0.11	0.28
Gesture.19	Point_NM_100	0.85	0.01	0.06	0.16