MAPPING NATURAL GESTURAL INPUTS TO TRADITIONAL TOUCHSCREEN INTERFACE DESIGNS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree of
Master of Science
Industrial Engineering

---

by
Katherina Jurewicz
December 2016

Accepted by:
David Neyens, PhD, MPH, Committee Chair
Kenneth Catchpole, PhD
Joel Greenstein, PhD
Scott Mason, PhD

ABSTRACT

Gestures are a natural means of every day human-human communication, and with the introduction of gestural input technology, there is an opportunity to investigate the application of gestures as a means of communicating with computers and other devices. The primary benefit of gestural input technology is that it facilitates a touchless interaction, so the ideal market demand for this technology is an environment where touch needs to be minimized. The perfect example of an environment that discourages touch are sterile environments, such as operating rooms. Healthcare-associated infections are a great burden to the healthcare system, and gestural input technology can decrease the number of surfaces, computers, and other devices that a healthcare provider comes in contact with. Gestural input technology has been investigated extensively in the operating room for surgeons manipulating radiological images but an application for anesthesia providers has not been investigated. The objective of this research was to map 3D gestural inputs to traditional touchscreen interface designs within the context of anesthesiology, and an experimental study was conducted to elicit intuitive gestures from users and assess the cognitive complexity of ten typical functions of anesthesia providers. Intuitive gestures were observed in six out of the ten functions without any cognitive complexity concerns. Two functions, of the remaining four, demonstrated a higher level gesture mapping with no cognitive complexity concerns. Overall, gestural input technology demonstrated promise for the ten typical functions of anesthesia providers in the operating room, and future research will continue investigating the application of gestural input technology for anesthesiology in the operating room.

DEDICATION

This thesis is dedicated to my family and friends, especially to my parents, Tom and Nedra Jurewicz, who provided me with everything I needed to be where I am today.

ACKNOWLEDGMENTS

There are many people who have helped me throughout my thesis work, and I am thankful for all of the support I have received. My advisor, Dr. David Neyens, has been essential in the progression and completion of my thesis. I could not ask for a better advisor as he always provided me with guidance and support while still allowing me the autonomy to shape my thesis into what I wanted it to be. I would like to thank my other thesis committee members for putting in the time and effort to ensure that the work I do is the best it can be. I would like to thank Dr. Kenneth Catchpole, Dr. Joel Greenstein, and Dr. Scott Mason for their valuable comments and guidance while serving on my committee and for encouraging me to do my best research.

I would like to again thank Dr. Neyens for hiring the best labmates to work with in the Ergonomics and Applied Statistics (EASt) Lab. Thank you, Ali, Courtney, Forrest, and Myrtede, for all of the help and encouragement you all provide me in the lab. I would like to especially thank Ali Hobbs and Forrest Sharp for the time they put in to validate the video data of gestures – I could not have completed my thesis without the work you put in for me.

Finally, I would like to thank my family and friends. Thank you to my mom and dad, my brother Peter, and my sister Tina, and all of my friends for believing in me and supporting me throughout my endless years of school and throughout life in general. I finally want to thank Dr. Ashley Childers who encouraged me to pursue a career in industrial engineering – I would not be here today without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1: INTRODUCTION

People commonly use gestures to communicate regardless of setting or spoken language. Gestures work fluidly with verbal cues and often support the clarification of thoughts and ideas. When used in conjunction with verbal communication, gestures give emphasis to speech, as in pointing in a specific direction while telling someone who is lost where to go. Gestures are also capable of replacing speech, especially in cases where verbal communication is not feasible (McNeill, 1992). For example, infants may use gestures with their parents signifying that they are hungry or travelers may use gestures to communicate with those who speak other languages. Regardless of the scenario, gestures demonstrate a natural way of communicating and are an integral part of human-human communication (Efron, 1941; Freedman, 1972; Kendon, 1988; McNeill, 1992). Since gestures are a part of everyday life, gestures are also capable of serving as a natural way to interact with computers and other devices (Karam & Schraefel, 2005). This human-computer gestural interaction differs from typical human-human gestural interaction, specifically in terms of gesture structure. Human-computer gestures can be static or dynamic, 2D or 3D, contact-based or vision-based, and emphasize body movement or hand movement. The more well-known human-computer gesture type is contact-based, 2D gestures. This would be the gestures that are used when interacting with a touchscreen, such as swiping the touchscreen on a phone to slide through pictures. The gesture type that is more analogous to natural human communication but less popular in everyday life is vision-based, 3D gestures. The 3D aspect of these gestures resonates with the 3D world that people live in, and the vision-based component gives a less intrusive feel as the person does not have to wear any sensors or other components (Baudel & Beaudouin-Lafon, 1993). Three-dimensional gestural input technology historically has had great success in the gaming industry with examples like Microsoft Kinect and Nintendo Wii. As these devices continue to develop and as new gestural input systems emerge, the opportunities of this technology go beyond the virtual gaming world.

Before extending gestural interaction to other applications, the foundation of gestures needs to be established. The gestures used to interact with computers and other devices can be grouped together to form a gestural language. Just as sounds and words form verbal languages, different movements and poses can be put together to create languages that bring meaning to movements. One of the challenges to creating a verbal language is organizing certain sounds, words, and phrases in a logical manner so that people can successfully convey a thought or idea. These sounds, words, and phrases are eventually categorized into parts of speech such as nouns, verbs, and adjectives. As gestural languages develop, there too is a need to logically organize particular movements and poses in order for someone to successfully convey a thought or idea through gestures. David Efron (1941) was the first fulfill this need in his pioneering work on the gestural behavior between Jewish and Italian immigrants and the people of New York City. Based off of Efron's (1941) classification system, other gesture classification systems have been proposed (Freedman, 1972; Kendon, 1988; McNeill, 1992). Regardless of the classification scheme or labeling of gesture types, there appears to be an agreement among four major categories that McNeill (1992) refers to as iconics, metaphorics, deictics, and beats. Iconics have a close relationship with speech and reiterate information given in speech, such as pointing up when saying "the spaceship went up!" (McNeill, 1992). Metaphoric gestures are similar to iconics in that the meaning from speech is reiterated but represent some abstract object such as pointing up when saying "my grades went up!" (McNeill, 1992). Diectics refer to gestures that point towards some indirect object, such as pointing at a computer screen to show where something is on the screen (McNeill, 1992). Beats refer to gestures that stress elements of speech such as a hand spreading as wide as it can to stress that something is big (McNeill, 1992). Kendon (1988) actually adds a fifth major category, symbolic, which would include those gestures that directly symbolize some inherent meaning; for example, sign languages would categorize as symbolic gestures. Sign languages are the epitome of a gestural languages as sign languages are full linguistic systems with different words, different phrases, comprehensive grammar structure, and they are understood by a community of users (McNeill, 1992); whereas, gestures in general do not have this sophisticated organization. These classification schemes represent the style of gestures in a universal manner, and they represent how gestures used in daily communication would evolve into categories based on their dependence on speech.

These taxonomies are only a means of describing and classifying gestures, which is actually just a preparatory step towards creating a gestural language and developing a gestural input system. In order to actually bring meaning to gestures, the first step in developing a gesture language is building a lexicon of gestures or a gesture vocabulary. There are a couple of approaches to defining these gesture languages: technology based and human based (Nielsen, Störring, Moeslund, & Granum, 2004). In the technology based vocabulary approach, gestures are defined by the capabilities of the technology where the primary goal is to maximize recognition accuracy (Nielsen et al., 2004). In this case, the gestures are taught to potential end users, and the functions of various applications are forced to work with these maximally recognized gestures (Nielsen et al., 2004). These systems are easy to implement but at the expense of usability (Nielsen et al., 2004). As for the human based approach, the focus is around usability. Instead of forcing the end user to use a set of gestures, the gestures are elicited through user studies. Both technology-driven methods (Baudel & Beaudouin-Lafon, 1993; Bizzotto et al., 2014; Freeman, Benko, Morris, & Wigdor, 2009; Mewes, Saalfeld, Riabikin, Skalej, & Hansen, 2016; Schroder, Loftfield, Langmann, Frank, & Reithmeier, 2014) and user-driven methods (Dong, Danesh, Figueroa, & El Saddik, 2015; Höysniemi, Hämäläinen, & Turkki, 2004; Jacob & Wachs, 2014; Jacob, Wachs, & Packer, 2013; Pereira, Wachs, Park, & Rempel, 2015; Stern, Wachs, & Edan, 2006; Wobbrock, Morris, & Wilson, 2009) are used in practice. In terms of which approach is superior, Morris, Wobbrock, and Wilson (2010) conducted a study to compare a gesture set elicited through end users to a gesture set developed by HCI researchers. The results demonstrated that participants preferred user-defined gestures over researcher-defined gestures suggesting that participatory design methodologies is critical when developing a gesture vocabulary (Morris et al., 2010).

Regardless of whether gesture lexicons are driven from technology or elicited from users, there are still several challenges to developing a successful gestural input system. As the quantity of gestures and functions grow within a gesture vocabulary set for a particular application, numerous concerns emerge with respect to both the technical and the human factors implications. Not only are gestural input systems more difficult to learn when there are more gestures in a gesture vocabulary set (Anderson & Bischof, 2013; Ardito, Costabile, & Jetter, 2014), there is a tradeoff between the number of gestures and the performance of the gestural input system (Wachs, Kölsch, Stern, & Edan, 2011). Furthermore, as these gesture vocabulary sets grow, there is a segmentation issue for the continuous capture of the gestures (Baudel & Beaudouin-Lafon, 1993; Pickering, Burnham, & Richardson, 2007). For example, a common gesture in gestural vocabulary sets is the open palm hand. If there is a dynamic gesture that incorporates movement of the open palm hand with a closing of the hand into a fist, the gestural input system needs to be capable of segmenting the movements to understand which gesture has actually been performed. There is also an occlusion problem of vision-based gestural input systems because the cameras rely on a visual of the hand and fingers, and if a person or an object occludes the camera, the gesture cannot be captured (Rautaray & Agrawal, 2015).

Aside from these more technical concerns, gestural communication can cause fatigue in the hand, wrist, and arms as they require more muscular effort than clicks on a mouse and keyboard (Baudel & Beaudouin-Lafon, 1993; O'Hara et al., 2014). In addition to this human factors implication, another important concern is ensuring that a gesture vocabulary set is appropriate for the context and application (Nielsen et al., 2004). Pereira, Wachs, Park and Rempel (2015) developed a gesture vocabulary set of

13 static and dynamic gestures to be used for 24 typical human-computer interaction functions. In this study, the issue of context sensitivity became apparent when multiple functions were assigned the same gesture (Pereira et al., 2015). When the number of functions outnumbers the number of gestures, gestural input systems need to be aware of the context. However, the challenge with gesture vocabulary sets that are developed for general human-computer interaction is that the context varies. When this is not done correctly and general gesture vocabulary sets are used for human-computer interaction across multiple contexts and applications, the system is not expected to succeed (Ardito et al., 2014). Gestural interfaces must consider the context in which it will be used and also incorporate new possibilities that the interaction could bring to that context (Wigdor & Wixon, 2011).

When taking the challenges currently known for gestural input systems, researchers can move forward in developing more effective and usable natural user interfaces. Historically, there are two approaches to developing new technologies: market pull and technology push. In technology push situations, the researcher starts with a new technology and matches it to an appropriate market (Ulrich & Eppinger, 2012). In market pull situations, the researcher starts with a market opportunity and finds a technology to meet the needs of the customer (Ulrich & Eppinger, 2012). In both situations there is concept development and design, but the main difference between the two is when the market demand is introduced in the design and development cycle. Market demand is either introduced as a question after a technology has been developed, or it is introduced as a need which pushes further develop of a specific technology. In other words, the market pull approach takes a market demand and develops a technology to fit the market demand instead of taking a technology and forcing it to fit within a system. Since it has been shown that eliciting gestures from potential users outperforms gestures designed by researchers (Morris et al., 2010), thus leaving opportunity for further technology development around end users, the market pull approach makes more sense for design and development of gestural input technology because the market pull approach takes a market demand from end users to push further development of a technology. In order to make the most of gestural input technology, a market demand needs to be identified and the gestural input technology needs to be assessed and further developed to meet the market demand effectively.

One of the benefits of gestural input technology is the fact that it facilitates a touchless interaction, so the ideal market demand for gestural input technology would be a situation where touch needs to be minimized. There are various situations where touch is discouraged, but the textbook case for discouraging touch would be sterile environments, such as an operating room (OR) (Wachs et al., 2008). Gestural input systems are the ideal technology for an OR environment because these are sterile environments that discourage excessive contact (Wachs et al., 2008). These systems have been implemented extensively in the OR for manipulation of radiological images during a surgical case (Bizzotto et al., 2014; Jacob & Wachs, 2014; Jacob et al., 2013; Mewes et al., 2016; O'Hara et al., 2014). The reason these systems were implemented was to avoid disrupting the surgical procedure and conserve sterility by allowing surgeons to interact with computers and other devices through gestures. Although much of the focus has been on surgeons and the ability to manipulate radiological images during a case, anesthesiologists also exhibit a demand for gestural input technology because they interact greatly with the patient before, during and after a surgical case (Miller & Pardo, 2011), but unlike the surgeon, they are not required to take additional measures (e.g., scrubbing in) to prevent patients from contracting an HAI.

In order to definitively determine if there a demand for gestural input technology for anesthesiology, it is necessary to understand HAIs overall and whether there is an opportunity to contract HAIs in the OR through an anesthesia provider. According to the U.S Department of Health and Human services (2013), HAIs can be contracted anywhere across the continuum of care for a patient, including the operating room (OR). In 2002, there were approximately 1.7 million cases of HAIs among U.S. Hospitals with 99,000 associated deaths (Klevens et al., 2007). In addition to the sheer quantity of HAIs, the cost burden is equally alarming. It is estimated that hospital-contracted HAIs account for $28 billion to $33 billion in healthcare costs every year (U.S. Department of Health and Human Services, 2013). As modern healthcare continues to increase in complexity, it is important to understand the etiology behind HAIs, determine the patterns of transmission for HAIs, and attempt to eliminate the risk of infection.

In 2010, The Society of Healthcare Epidemiology of America (SHEA) offered a national approach to HAIs (The Research Committee of the Society of Healthcare Epidemiology of America, 2010). Since its release, numerous studies have expanded the understanding of HAIs and proposed prevention techniques to be implemented hospital-wide, including but not limited to further training (Barsuk, Cohen, Feinglass, McGaghie, & Wayne, 2009; Comer et al., 2011), improvement in hand hygiene (Pittet et al., 2000; Sax et al., 2007), and best practices guidelines for healthcare providers (Marschall et al., 2014). As part of the hospital, the OR also incorporates these techniques, but because of the nature of work in the OR (i.e., interaction with one patient over a long period of time), these measures may not be enough to eliminate contamination (Stackhouse et al., 2011). With respect to the healthcare environment, the surface environment has been extensively connected to HAIs (Weber, Anderson, & Rutala, 2013); pathogens can survive on hospital room surfaces and medical equipment for hours, days, and even up to months (Weber et al., 2013). As healthcare providers, including the anesthesia team, care for multiple patients while touching these surfaces and equipment, they are facilitating the transfer of bacteria from one patient to another.

Regarding anesthesiology specifically, studies have begun quantifying the infection control issue for anesthesia providers. As for hand hygiene as a whole, Biddle and Shah (2012) observed an average of 34-41 hand hygiene opportunities per hour, and of these opportunities, 82% of the time anesthesia providers did not comply to hand hygiene practices during the perioperative period. With such a high noncompliance rate to hand hygiene, it is important to understand the behavioral patterns of anesthesia providers during the perioperative period and the potential role of the provider in the spread of bacteria. It was shown that anesthesia providers completed only 13 hand decontamination events while touching 1,132 objects during an observation period of 8 hours, with the anesthesia machine and anesthesia keyboard having the highest number of touches (Munoz-Price et al., 2013). In a separate study, the patient bed was shown to be the most frequently touched object with a mean of 77 touches per hour (Rowlands et al., 2014). When just looking at the intubation process, researchers were able to understand the dynamics of bacterial transmission by using fluorescent marker to represent contamination (Birnbach, Rosen, Fitzpatrick, Carling, & Munoz-Price, 2015). Although the fluorescent marker was initially present only inside the mouth and on the lips of the patient simulator, the fluorescent marker spread throughout the anesthesia environment during the intubation process (Birnbach, Rosen, Fitzpatrick, Carling, & Munoz-Price, 2015). Thirteen areas within the anesthesia environment (including the IV hub, anesthesia machine surface, anesthesia circuit, oxygen valve, and anesthesia cart) were contaminated in 100% of the experimental sessions, and the computer keyboard was contaminated in 80% of the experimental sessions (Birnbach, Rosen, Fitzpatrick, Carling, & Munoz-Price, 2015). This study demonstrates that there is potential for widespread bacteria contamination before the operation even begins. A separate study showed that during the operation, the anesthesia environment was involved with bacterial transmission in 89% of the observed cases (Loftus et al., 2011). These findings support the notion that there is a cyclical pattern of bacterial transmission from the patient to the anesthesia environment back to the patient. This pattern supports corresponding research that shows the anesthesia providers' contaminated hands play a key role in bacterial transfer (Loftus et al., 2012). This a major concern for infection control because patients are at risk of being infected with their own bacteria, and since not all of the bacteria on surfaces and objects are completely removed, future patients are at risk of being infected by the bacteria that is immediately present on surfaces and objects within the anesthesia environment (Stackhouse et al., 2011).

Researchers understand this infection control problem and have sought to improve contamination in the anesthesia work environment. One study found that setting up the anesthesia environment to keep clean and dirty work areas separated reduces the amount of contamination from the start to the end of a surgical case (Clark, Taenzer, Charette, & Whitty, 2014). Furthermore, there have also been developments related to changes in work practices. With regards to the practice of wearing two pairs of gloves for intubation and taking one pair off after completing intubation, the number of contaminated areas within the anesthesia workstation is reduced when compared to the standard practice of wearing a single pair of gloves (Birnbach, Rosen, Fitzpatrick, Carling, Arheart, et al., 2015). With respect to hand hygiene, Koff et al. (2009) improved the proximity of a hand hygiene system by having the anesthesia provider use a body-worn hand sanitation device. When using this device in addition to having a wall-mounted dispenser and dispenser on the anesthesia cart, the number of hand decontamination events increases and contamination of the anesthesia machine decreases (Koff et al., 2009). However, Koff et al. (2016) eventually demonstrated that when using this device and given hand decontamination event feedback, there was not a reduction in 30-day postoperative HAIs. In other words, although this system improved hand hygiene and contamination in the OR, there was no association between use of the device and a reduction in HAIs (Koff et al., 2016). While these interventions demonstrate some positive change, the amount of research being done in this area does not match the evident infection control concern in the anesthesia environment.

Anesthesiology, health technology, and healthcare in general will continue to grow in complexity, and as this occurs, it is crucial to reduce and ultimately eliminate the risk of infection in the OR. The anesthesia environment and the anesthesia provider play key roles in the transmission of bacteria during the perioperative care of a patient. If anesthesia providers can reduce the number of surfaces and objects they come in contact within the anesthesia environment, there can be a potential reduction in risk of infection to the patient. A completely touchless OR would be the ideal in relation to sterility, but this is not possible with the current work practices of anesthesiology. Although a touchless OR could be futuristic, the technology currently exists to facilitate a number of touchless interactions through gestural communication.

The market demand is clear for anesthesiology in the OR as infection control has shown to be an issue, and there is an opportunity to determine if gestural input technology makes sense as an intervention for anesthesia providers in the OR to improve infection control numbers. In order to do so successfully, gestures should be elicited from users (Morris et al., 2010) and be suitable for the context and domain in which it is applied (Ardito et al., 2014; Nielsen et al., 2004; Wigdor & Wixon, 2011). Although gestural input systems have previously been developed for the OR, anesthesiology cannot simply adopt these systems since the context and work domain of anesthesiology is considerably different. Additionally, in these systems the gestures act as a navigational tool for manipulating or rotating an image, but they are not command specific or function specific (e.g., silencing an alarm) and thus do not fully incorporate all of the capabilities of gestural input technology. There is an opportunity for anesthesiology to learn from previous research in gestural interface designs to develop an effective vision-based, 3D gestural input system to help minimize the infection control issue in anesthesiology.

Research Objective

Touchscreen displays, and consequently 2D gestures, are already common in the OR (Hurka, Wenger, Heininger, & Lueth, 2011), but there is not a clear mapping of 3D gestures to functions of these displays. The application of 3D gestures in anesthesiology would reduce the need for touching multiple surfaces in the anesthesia work environment. The overall objective of this research is to map 3D gestural inputs to functions of traditional touchscreen interface designs in the context of anesthesiology. The first aim of this research is to identify the gesture-function mappings that are most intuitive to the user (Aim 1). The second aim of this research is to determine the cognitive complexity associated with each gestural-function mapping (Aim 2).

CHAPTER 2. METHODOLOGY

The previous chapter provided the background necessary for understanding this research by discussing gestural input technology and the market demand anesthesiology possesses in the OR. This chapter will outline the methodology used to evaluate the potential use of gestural input technology in the OR for anesthesiology, specifically the intuitive gesture-function mappings and the cognitive complexity of gesture-function mappings. First an overview of the participants, apparatus and setting, study design, independent and dependent variables, and the overall procedure of the experiment is presented. A description of the data analysis methodology for the respective aims is also given.

Participants

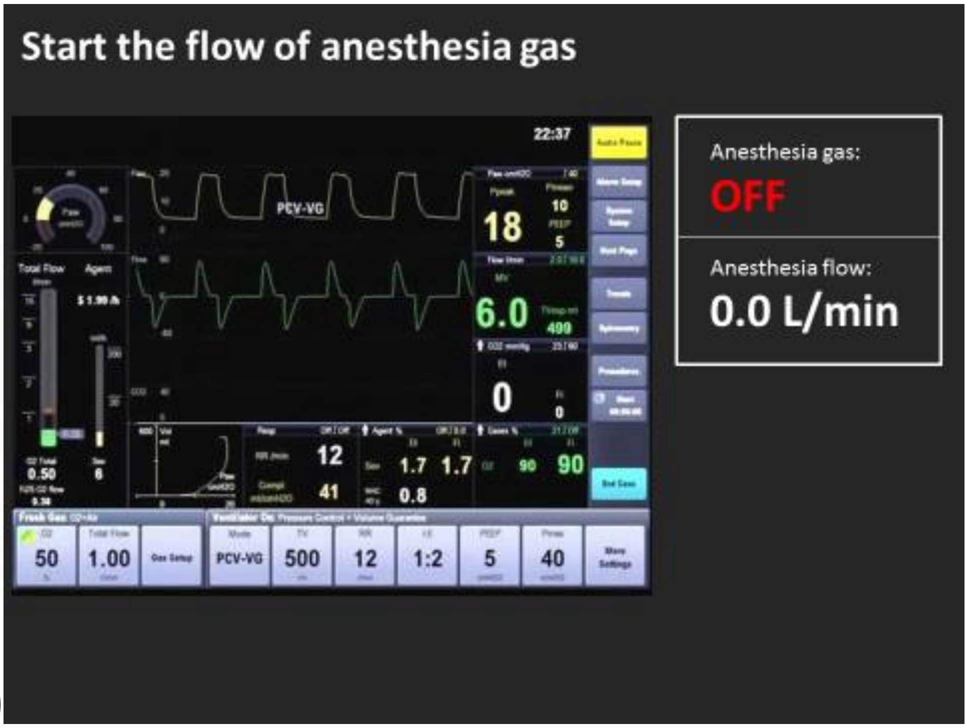
Participants (N=30) were required to be able to read, write, and speak in English and had full manual dexterity of fingers, wrists, and arms in their non-dominant hand. Participants were recruited from Clemson, SC and surrounding areas and received a compensation of $10 for one hour of their time. The study was approved by the Clemson University IRB (IRB: 2016-110).

Apparatus and Experimental Setting

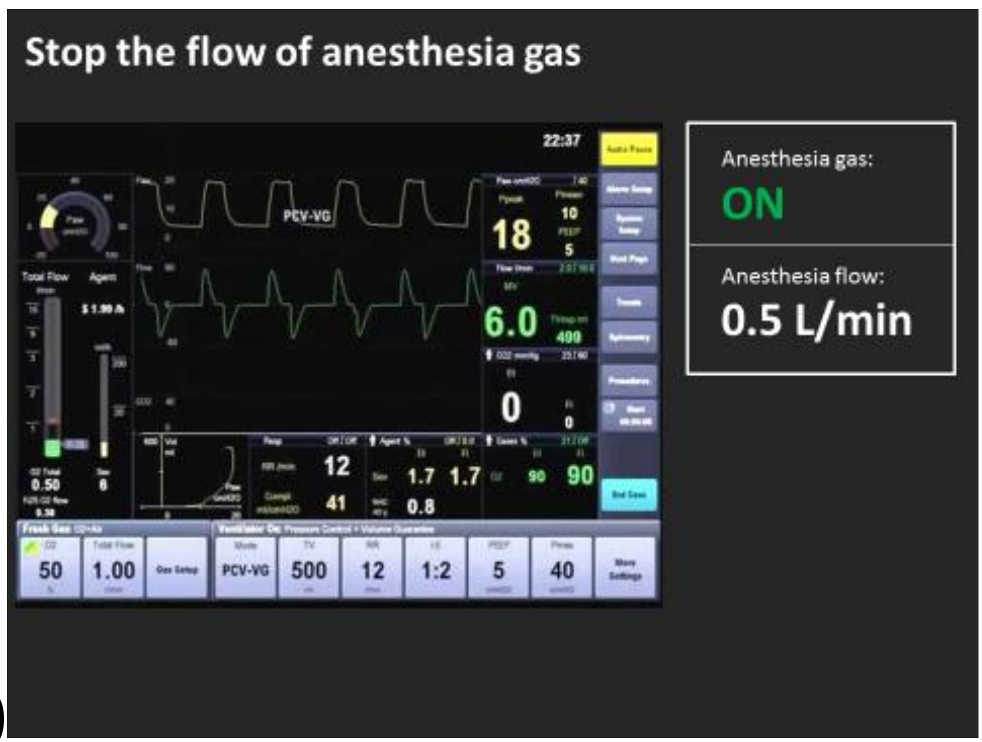
This study used an Intel RealSense Camera which is capable of detecting 3D gestures. The camera was mounted to a 22- inch Dell desktop monitor. Since none of the participants were anesthesia providers, it was important for this study to replicate an anesthesia setting in an OR. In order to do so, the sound of the pulse oximeter was played throughout the experimental session, and when alarms were relevant to the experiment, the alarm of a patient whose heart shows no electrical activity sounded. Participants were required to wear a non-latex glove just as an anesthesia provider would in the OR. Wearing the non-latex glove aids in replicating the environment in which an anesthesiologist works because healthcare providers are recommended to wear gloves when working with a patient (World Health Organization, 2009). The participant used their non-dominant hand to interact with the camera in order to imitate a situation where an anesthesia provider is working with a patient and their dominant hand is occupied. The experimental setting is shown in Figure 1.

Figure 1. Experimental setting

The functions of the experiment reflected the role of an anesthesia provider. These functions were chosen based off of in-person and video observations of anesthesia personnel in an OR. Additionally, Wigdor & Wixon (2011) recommend to test new functionalities that gestural input technology will incorporate, so there was one function that reflected a new functionality of a gestural input system. These functions are described in Figures 2, 3, 4, 5, and 6. Function 9 shown in Figure 6a is a new function that would be incorporated if a gestural input system were introduced in anesthesiology. Unlike the other functions where the activity is currently performed within the context of anesthesia, Function 9 is a functionality that is part of a new gestural display that would be added to the anesthesia context and applicationEach function was presented on a separate slide within a Microsoft PowerPoint presentation. The experimental setting adopted a Wizard of Oz technique in which the experimenter acts as a "wizard" and simulates the behavior of a complete system. In other words, the gesture was perceived to interact with the computer, but the experimenter manually advanced to the next slide after completion of a gesture. Wizard of Oz methodologies have been shown to be successful in user-elicitation studies with gestures (Aigner et al., 2012; Freeman et al., 2009; Höysniemi et al., 2004; Morris et al., 2010) since the wizard is quicker to interpret gestures and manually progress throughout an experiment than the gestural input technology, which removes any potential frustration a participant may have with the system. Additionally, since this study was a user based approach, the gestural input system has not fully been developed for a specific application yet as this research focused on eliciting intuitive gestures from users and investigating cognitive complexity so that a full gestural input system could eventually be developed. In a conventional Wizard of Oz study, the wizard is out of the room and unseen, but in this study, the wizard was in the room advancing through the slides.

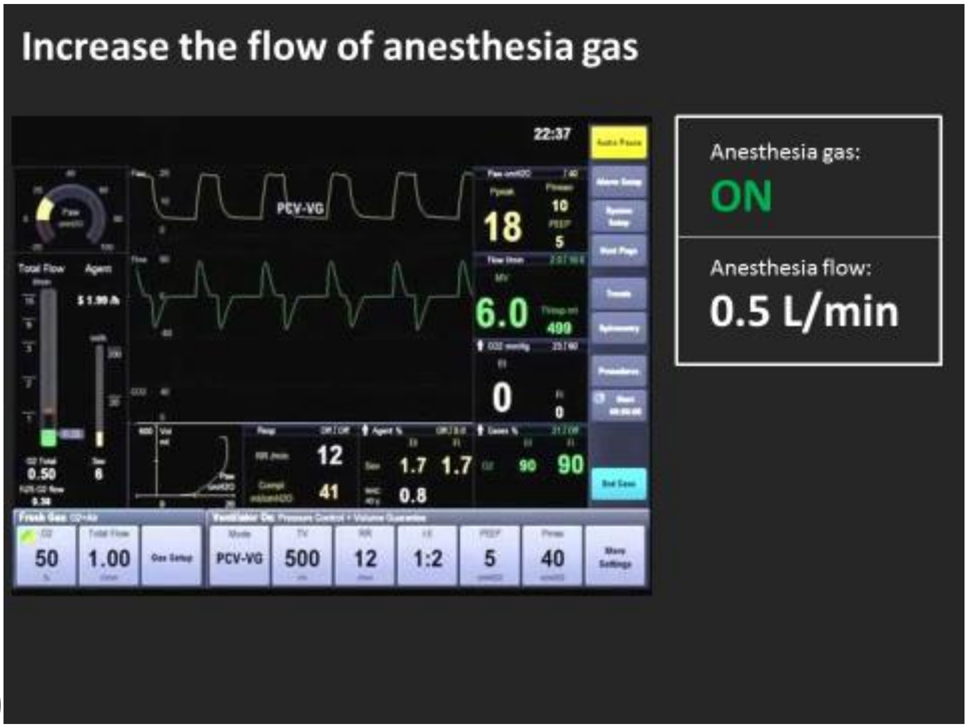


a)

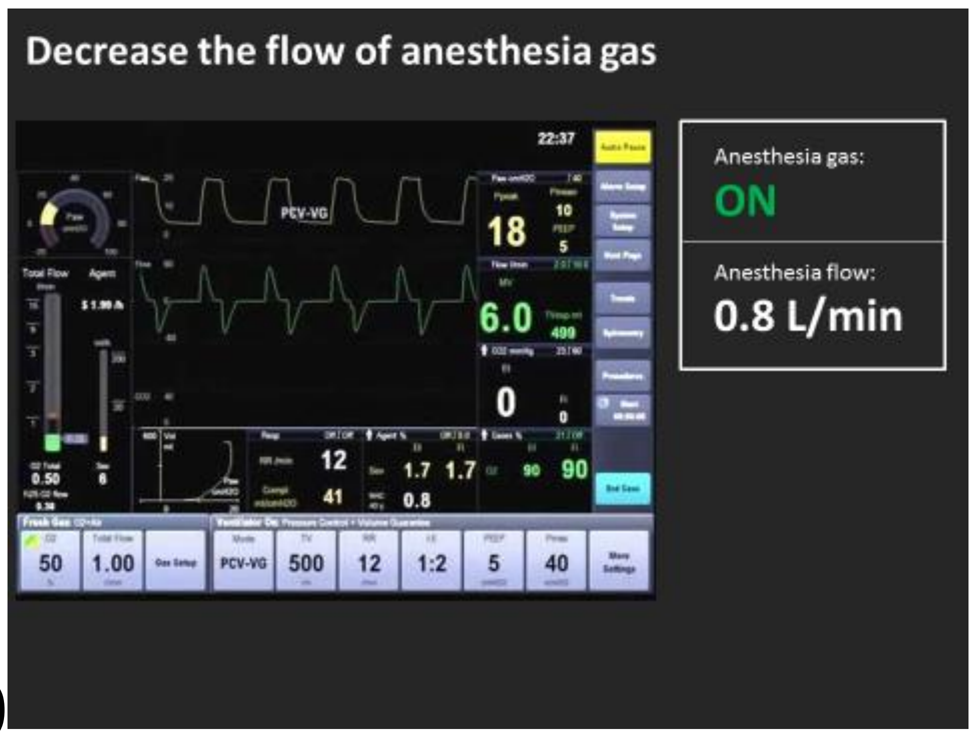


b)

Figure 2. Functions for gesture mapping. (a) Function 1 – Start the flow of anesthesia gas. (b) Function 2 – Stop the flow of anesthesia gas

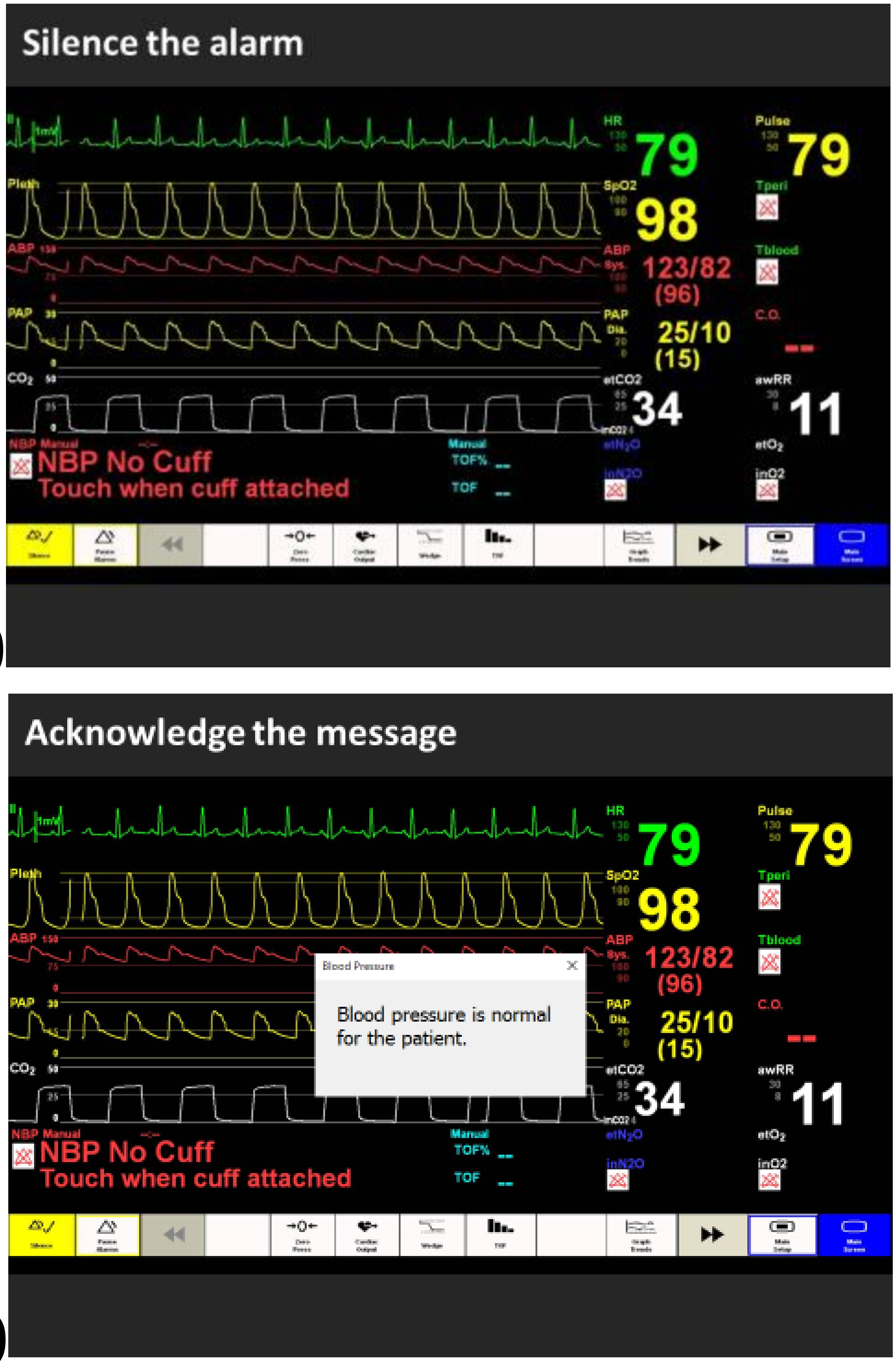


a)



b)

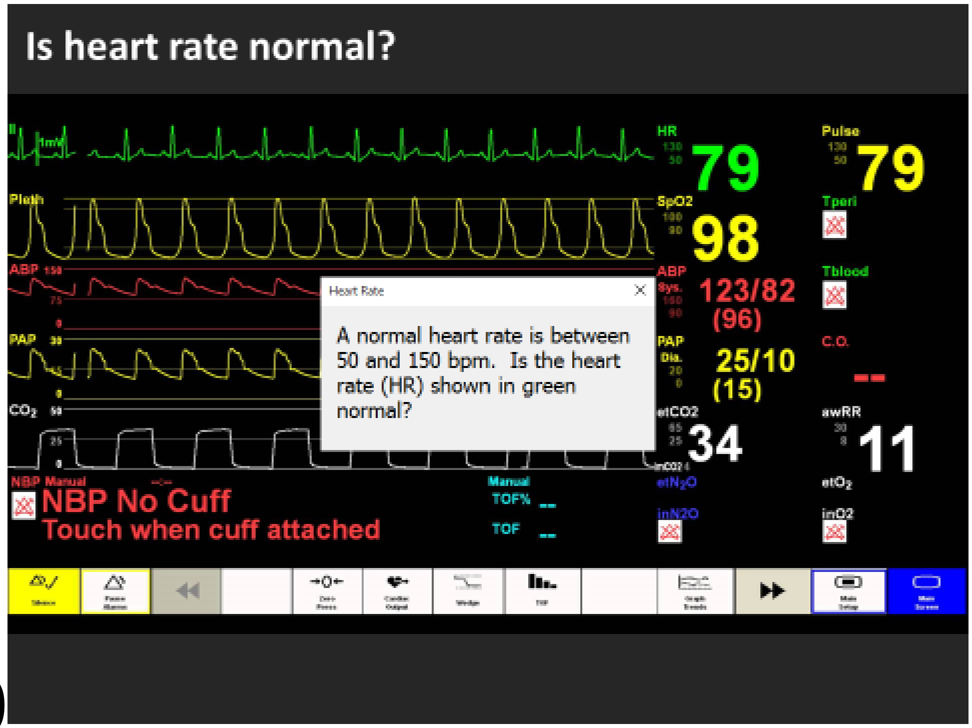
Figure 3. Functions for gesture mapping. (a) Function 3 – Increase the flow of anesthesia gas. (b) Function 4 – Decrease the flow of anesthesia gas



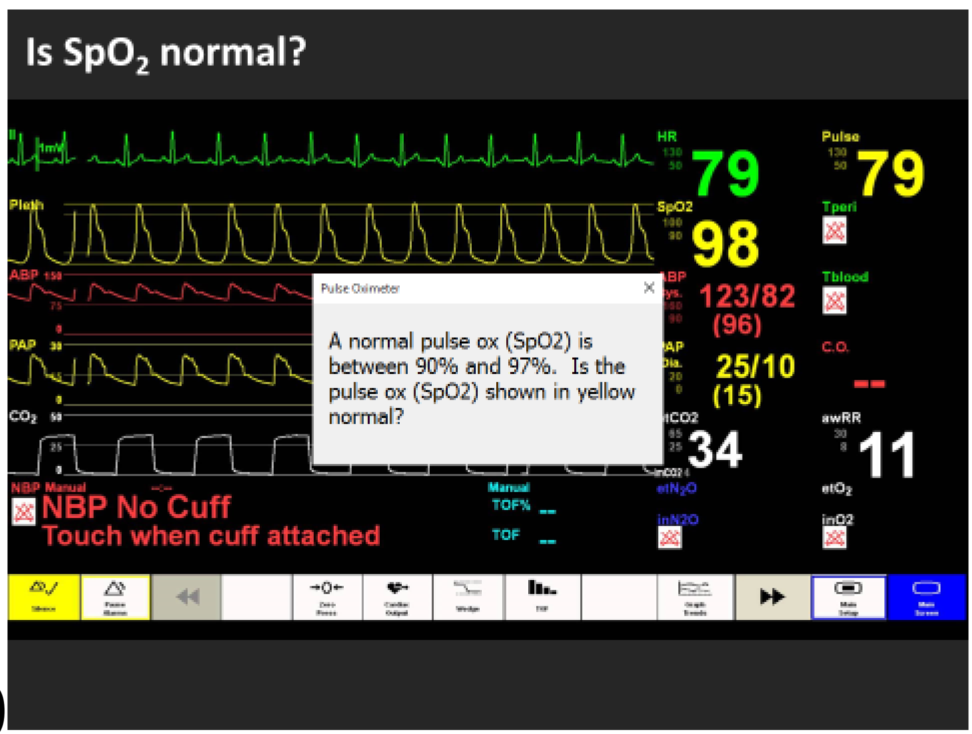
a)

b)

Figure 4. Functions for gesture mapping. (a) Function 5 – Silence the alarm. (b) Function 6 – Acknowledge the message



a)



b)

Figure 5. Functions for gesture mapping. (a) Function 7 – Is heart rate normal? (b) Function 8 – Is SpO_2 normal?

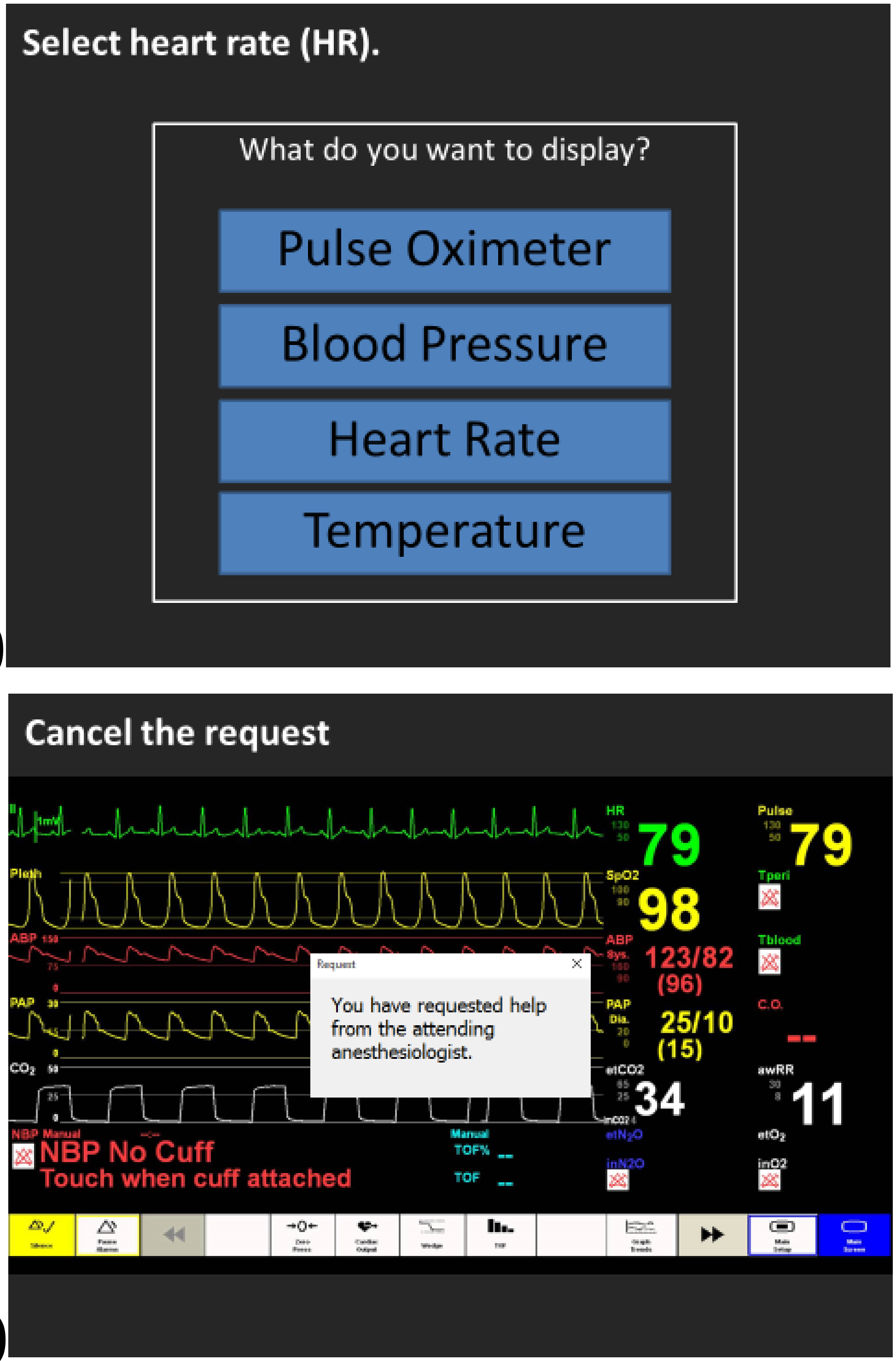


Figure 6. Functions for gesture mapping. (a) Function 9 – Select heart rate (HR). (b) Function 10 – Cancel the request

Study Design

The experiment was a repeated measures design where the repeated element was the function (N=10). The 10 functions were presented in random order within a block and then randomized for two additional blocks, so that the participant chose gestures for each function three separate times. This yielded 30 total gestures for each participant. The study design with respective aims are shown in Figure 7.

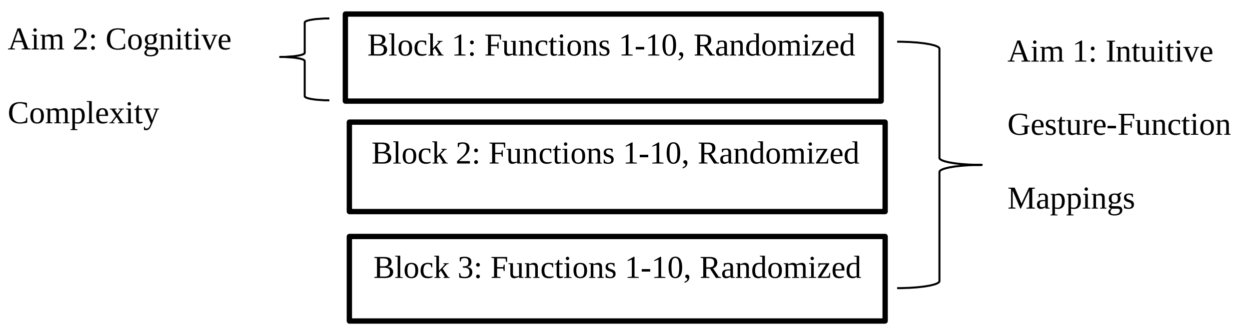


Figure 7. Study design with respective research aims

Experimental Procedure

Once the participant arrived, the experimenter verbally reviewed the Informed Consent Form and obtained the participant's written consent. All participants received a copy of this form at the end of their visit. The participant was then asked to complete some questionnaires: Complacency-Potential Rating Scales on the reliability of systems and a demographic survey. This study adopted and modified Nielson et al.'s (2004) human-centered approach for developing intuitive and ergonomic gestural interfaces. As part of this approach, the participant practiced a set of gestures to familiarize themselves with gestural input technology. The researcher first asked the participant to put a non-latex glove on their non-dominant hand. After doing so, the researcher reviewed the set of gestures that they practiced. The participant performed a gesture 15 times, and this was repeated for each of the gestures provided. The gestures that were practiced were the ones defined in the Intel RealSense Software Development Kit (Appendix A).

As part of Nielsen et al.'s (2004) bottom-up approach, the effect (i.e., the function) was given and the user was to perform the cause (i.e., the gesture) in the experimental session. For example, the participant was asked to "Silence the alarm" by performing the gesture that created that effect. The participant was told to use any gesture that they want and whichever gesture was their "first guess" to create the effect. After the participant completed the 30 total functions, the participant completed the Perceived Ease of Use portion of the User Acceptance Survey. The participant was then debriefed and given compensation.

A video of the participant's hands and fingers was recorded from the start of the practice session until the end of the experiment, and this video also recorded the computer system time. During the experiment, a JavaScript program recorded which gestures are captured by the camera as well as the computer system time. These data were saved to a text file. A Microsoft Visual Basic program recorded the computer system time in Microsoft Excel for each slide advancement during the experiment.

Independent Variables

The independent variables for this study were display function, age, gender, race, handedness, education (highest degree obtained), education (major and/or minor), video game use, and experience with virtual reality gaming.

Dependent Variables

The dependent variables for this study were gesture chosen and response time. The response time was defined at the duration of time from a gesture presented to gesture chosen. The gesture chosen supported investigation of Aim 1 of this research, and response time supported investigation of Aim 2.

Aim 1 Methodology

To identify the gesture-function mappings that are most intuitive to the user, the most frequent gesture-function mapping needed to be determined. The frequency of gesture-function mappings is an accurate indication of intuition (Nielsen et al., 2004). In order to classify gestures, a gesture dictionary was built to include the gestures practiced plus additional gestures used in other studies as well as commonly used gestures (Appendix B). For each function, the participant's chosen gesture was classified by three researchers independently via video analysis of the participant's hands and fingers. If there was a discrepancy between any of the researchers classifications, then all three researchers reviewed the video together to reach an agreement; thus, consensus building was used to determine the gestures chosen.

Before determining the most frequent gesture-function mappings across participants, the internal consistency of each participant must be predetermined. When participants are inconsistent in choosing gestures for a function, it demonstrates that there is a lack of internal intuition for that function. If there is a lack of internal intuition, this data should not be included when determining intuition across all participants. Thus, the first step in determining the gesture-function mappings that are most intuitive to the user was to determine the internal consistency for each participant. This was done by comparing the gesture chosen for a function for each participant across blocks. For example, if a participant chose the same gesture in all three iterations of a function, they were labelled as "Completely Internally Consistent" for that function. If a participant chose the same gesture for two iterations of a function and a different gesture for one iteration, they were labelled as "Partially Internally Consistent" for that function. If a participant chose three different gestures for all three iterations of a function, they were labelled as "Internally Inconsistent." The data for participants who were labelled as "Completely Internally Consistent" and "Partially Internally Consistent" were used to construct a table of gesture-function mappings. A tally was recorded of gestures chosen among all gestures for a function, and the gesture that occured most frequently is the gesture-function mapping that is deemed to be the most intuitive.

Aim 2 Methodology

Response time is often used to provide an indication of cognitive complexity (Horsky, Kaufman, Oppenheim, & Patel, 2003), so for this study, a longer response time for a gesture suggests that there is a higher cognitive load. Response times were determined by comparing the system time of the slide advancement to the system time shown in the video of the participant's hands and fingers. The responses of variable

“Video Game Use” were collapsed into two categories: “Yes” to playing video games and “No” to not playing video games. The “Yes” category included all positive responses to the video game use question from the demographic survey, and the “No” category included the negative response of “Do not play” video games. The responses to “What is your major/minor” were also collapsed into “Science and Engineering” and “Not Science and Engineering.” The “Science and Engineering” category included all participants who majored or minored in science or engineering, and the other category included all other majors. The variable “Computer Use” was not used in the analysis as all participants reported frequent computer use. A mixed effects linear regression model was developed to determine the factors which are associated with long response times (i.e., the factors that are associated with a high cognitive load). The equation of the mixed effects linear regression model in matrix notation is shown below:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$$

where:

$\boldsymbol{y}$ is an *N x 1* column vector of the response variable

$\boldsymbol{X}$ is an *N x p* matrix of *p* predictor variables

$\boldsymbol{\beta}$ is a *p x 1* column vector of the regression coefficients of the fixed effects

$\boldsymbol{Z}$ is an *N x q* matrix of *q* random effects

$\boldsymbol{\gamma}$ is a *q x 1 column* vector of the random effects

$\boldsymbol{\varepsilon}$ is an *N x 1* column vector of the residuals

The variables included in the regression model are shown in Table 1. The random effect in this regression model is the participant, X_8; therefore, the intercept of the linear regression model was used to adjust for differences between participants. All other variables were treated as fixed effects in the linear regression model. Only data from Block 1 were used in the analysis so as to investigate a function's response time at the first time a function is presented to the participant.

Table 1. Summary of Response and Predictor Variables

Variable	Name	Variable Type
Y	Response time	Continuous
F_i	Function	Categorical
X_1	Age	Continuous
X_2	Gender	Categorical
X_3	Handedness	Categorical
X_4	Education, highest degree obtained	Categorical
X_5	Education, area of study	Categorical
X_6	Video game use	Categorical
X_7	Virtual reality gaming experience	Categorical
X_8	Participant ID	Categorical

Before fitting this mixed effects model, an ANOVA was performed to compare two linear models: a linear model with a fixed intercept plus the random effect and a null model with only the fixed intercept. If the P-value is <0.001, then the mixed model was preferred over the null model. After fitting a model, diagnostic tests were performed to ensure the assumptions for the linear model is met: linearity, homescedacity, normality, independence, and no multicollinearity issues. To identify any multicollinearity issues, VIF values were calculated and any predictor variables with VIF values >5 were removed from the model. Any influential points were also removed from the data set by calculating Cook's distance. Cook's distance is a measure for one unit's influence on parameter estimates (Cook, 1977). The formula for calculating Cook's distance is shown below:

$$\mathrm{D}_i = \frac{{e_i}^2}{s^2 p}\left[\frac{h_i}{(1-h_i)^2}\right]$$

where:

D_i is Cook's distance for the *i*th observation

e_i is the residual for the *i*th observation

s^2 is the mean squared error of the regression model

h_i is the leverage of the *i*th observation

For mixed models, a point is regarded as influential if the respective Cook's Distance value exceeds the cut off value of (Van der Meer, Te Grotenhuis, & Pelzer, 2010):

$$4/n$$

where *n* refers to the number of groups of the grouping variable.

The mixed effects linear regression model can only determine if functions are associated with response times compared to one reference function, so in order to compute differences in response times for each pair of functions, Tukey contrasts were calculated to make the pairwise comparisons. R version 3.3.2 was used to do the analysis and used the *skewness* function of the e1071 package (Meyer, Dimitriadou, Hornik, Weingessel, & Leisch, n.d.), the *lmer* function of the lme4 package (Bates, Mächler, Bolker, & Walker, 2014), the *glht* function of the multcomp package (Hothorn, Bretz, & Westfall, 2008), and the *cooks.distance* function of the influence.ME package (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012).

CHAPTER 3. RESULTS

The previous chapter discussed the methodology used to obtain the results that will be discussed in this chapter. In this chapter the results will be broken down into four sections –overview of study participants, intuitive gesture-function mappings, cognitive complexity of gesture-function mappings, and general findings.

Overview

The characteristics of the participants for this study are described in Table 2. The mean response time across all blocks and participants was 4.77 seconds, and the standard deviation for response time was 2.93 seconds. The variable "Computer Use" was not analyzed since every participant responded that they used a computer.

Table 2. Characteristics of study participants

Variable Name	N (%)
Age	M=21.80, SD=2.23
Gender	
Male	15 (50)
Female	15 (50)
Handedness	
Right	26 (86.7)
Left	3 (10)
Ambidextrous	1 (3.3)
Education, highest degree obtained	
High School/GED	16 (53.3)
Bachelors	11 (36.7)
Masters	3 (10)
Education, area of study	
Science or Engineering	19 (63.3)
Not Science or Engineering	11 (26.7)
Video Game Use	
Yes	15 (50)
No	15 (50)
Virtual Reality Gaming Experience	
Yes	14 (46.7)
No	16 (43.3)

Aim 1 Results

Internal Consistency

Recall from Chapter 2 that participants who exhibited internal inconsistency for a function had their data removed for that particular function in the gesture mapping analysis. A summary of the internal consistencies is in Table 3. As shown in the last column in the table, no more than 3 participants had data removed for a particular function.

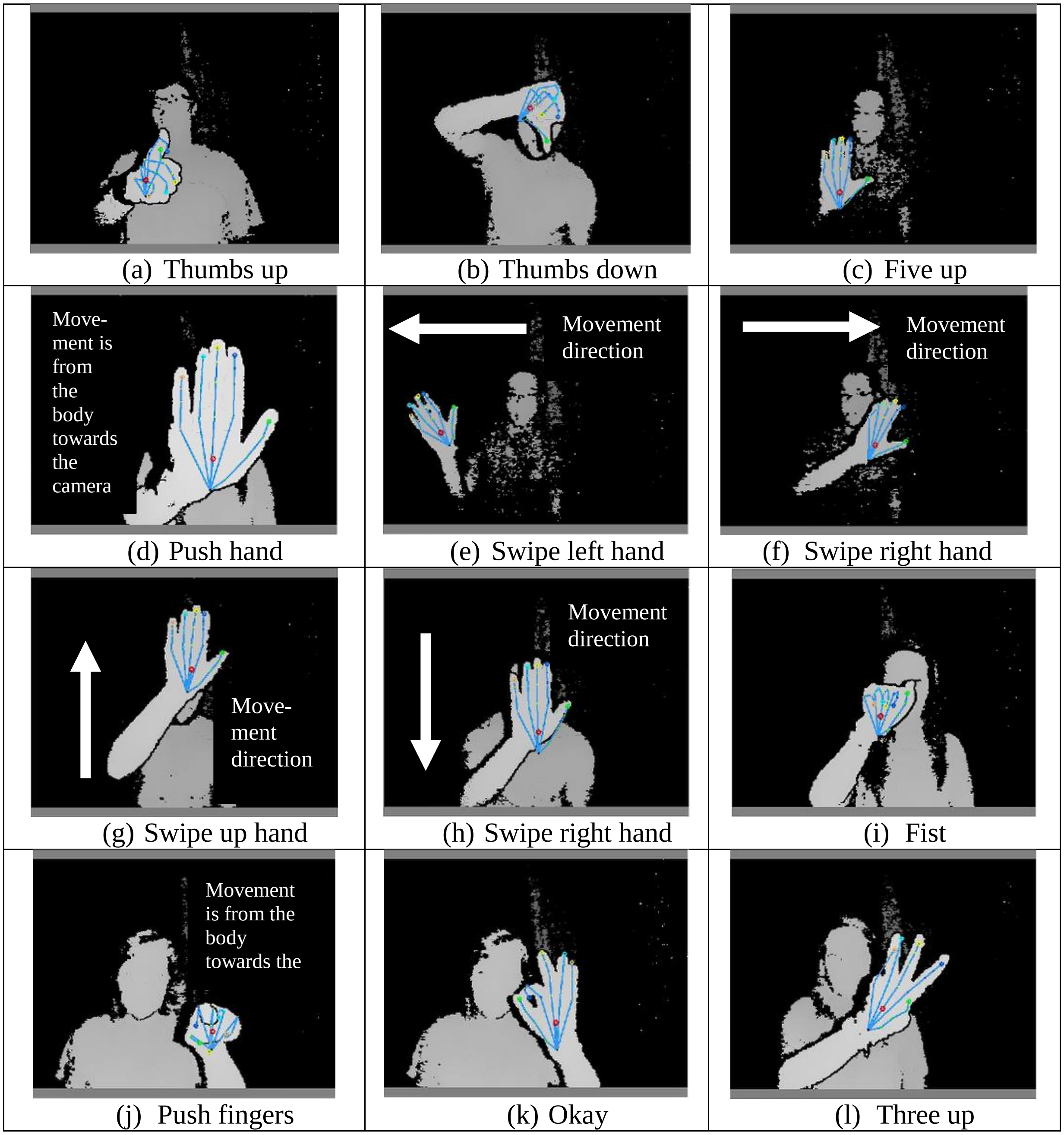
Table 3. Internal consistencies of all participants (N=30)

Function	Degree of Internal Consistency		
	Complete	Partial	Inconsistent
1	16	12	2
2	17	10	3
3	21	9	0
4	19	11	0
5	14	14	2
6	12	15	3
7	21	6	3
8	13	14	3
9	20	10	0
10	15	14	1

Note: Inconsistent participants were excluded in gesture-function mapping analysis

Gesture-Function Mappings

The participants who had some degree of internal consistency were used to determine the most frequent gesture chosen for each function. Overall, 42 unique gestures were performed across participants and across functions, and a total of 852 gestures were taken into account for the analysis. Among these unique gestures, a pictorial representation of the most common gestures is provided in Figure 8.



(a) Thumbs up
(b) Thumbs down
(c) Five up
(d) Push hand
(e) Swipe left hand
(f) Swipe right hand
(g) Swipe up hand
(h) Swipe right hand
(i) Fist
(j) Push fingers
(k) Okay
(l) Three up

Figure 8. Pictorial representation of commonly used gestures

Function 1 (see Figure 2a) and Function 2 (see Figure 2b) are opposites, "Start the flow of anesthesia gas" and "Stop the flow of anesthesia gas." The histograms for gestures chosen for functions 1 and 2 are shown in Figures 9 and 10 respectively. For Function 1 (Start the flow of anesthesia gas), the most frequently chosen gesture was "thumbs up" (Figure 8a) and was performed a total of 29 times. The second most frequently peformed gesture was "swipe up hand" (Figure 8g) but was only performed 14 times. The rest of the gestures for function 1 were performed less than 10 times. For function 2 (Stop the flow of anesthesia gas), the most frequently chosen gesture was "five up" (Figure 8c) and was performed 17 times. The second most frequently chosen gesture was "push hand" (Figure 8d) and was performed 14 times. The third and fourth most frequently chosen gesture was "fist" (Figure 8i) and "swipe left hand" (Figure 8e) and were performed 12 and 10 times, respectively. For function 2, there was a not a substantial gap between the top four chosen gestures; although when looking at the top two chosen gestures (five up and push hand), "push hand" (Figure 8d) is actually a dynamic movement of the "five up" (Figure 8c) static gesture.

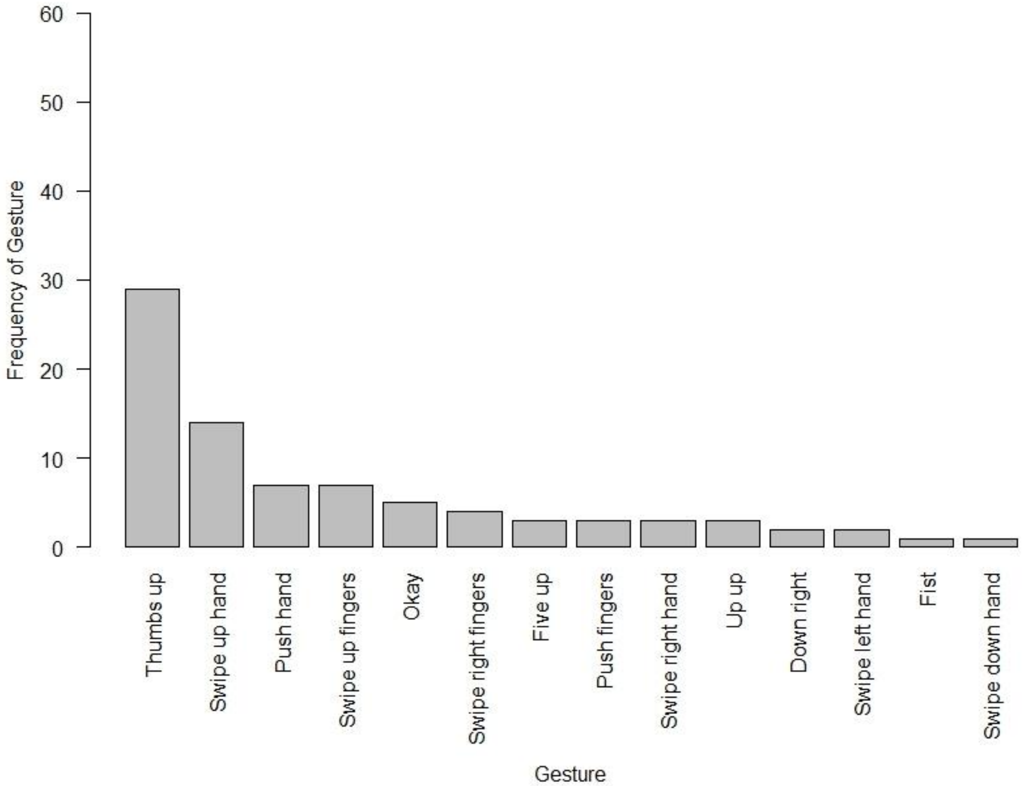


Figure 9. Histogram of gestures chosen for Function 1: Start the flow of anesthesia gas

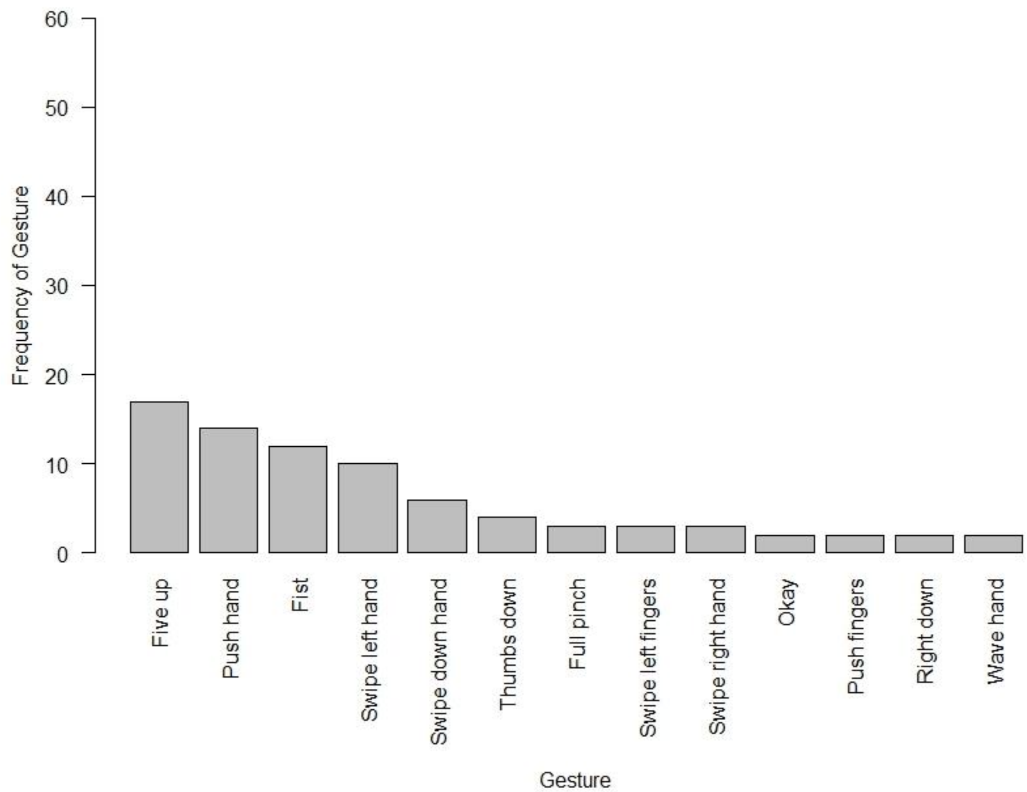


Figure 10. Histogram of gestures chosen for Function 2: Stop the flow of anesthesia gas

Similar to function 1 and function 2, function 3 (see Figure 3a) and 4 (see Figure 3b) are also polar opposites. The histograms for gestures chosen for these two functions are shown in Figures 11 and 12, respectively. For function 3 (Increase the flow of anesthesia gas), the most frequently chosen gesture was "swipe up hand" (Figure 8g) and was performed a total of 33 times. The second most frequently performed gesture was "thumbs up" (Figure 8a) but was only performed 16 times. The rest of the gestures for function 3 were performed 10 or less times. For function 4 (Decrease the flow of anesthesia gas), the most frequently chosen gesture was "swipe down hand" (Figure 8h) and was performed a total of 33 times. The second most frequently performed gesture was "thumbs down" (Figure 8b) but was only performed 17 times. The rest of the gestures for function 4 were performed less than 10 times.

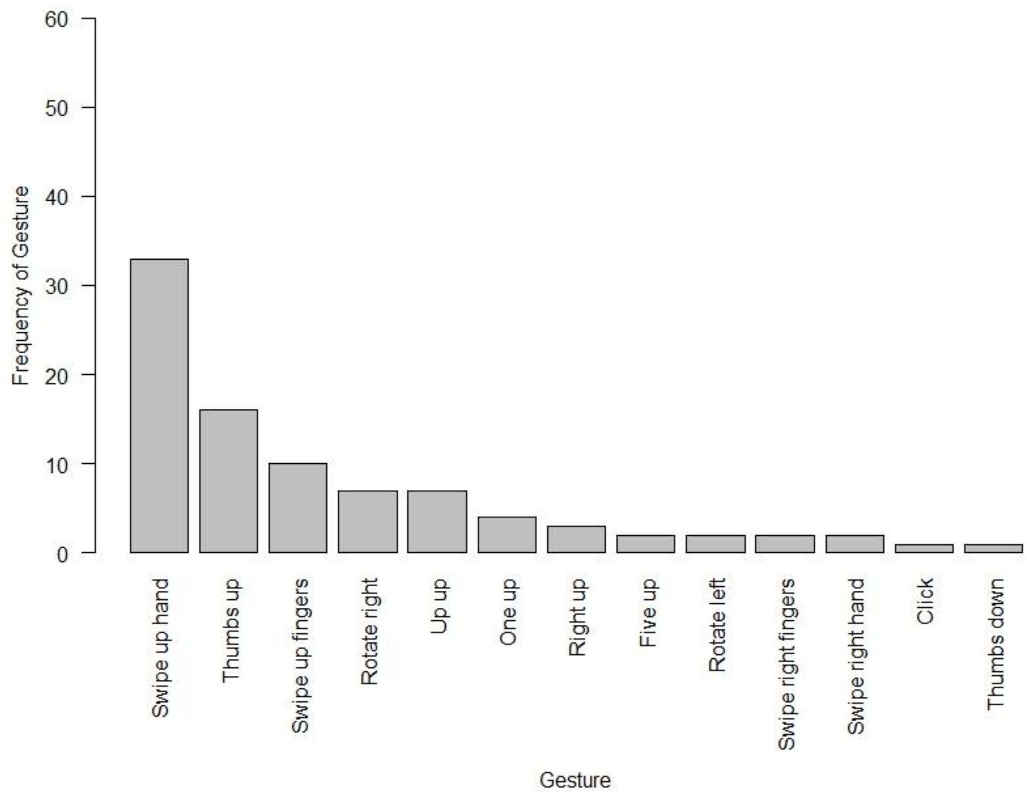


Figure 11. Histogram of gestures chosen for Function 3: Increase the flow of anesthesia gas

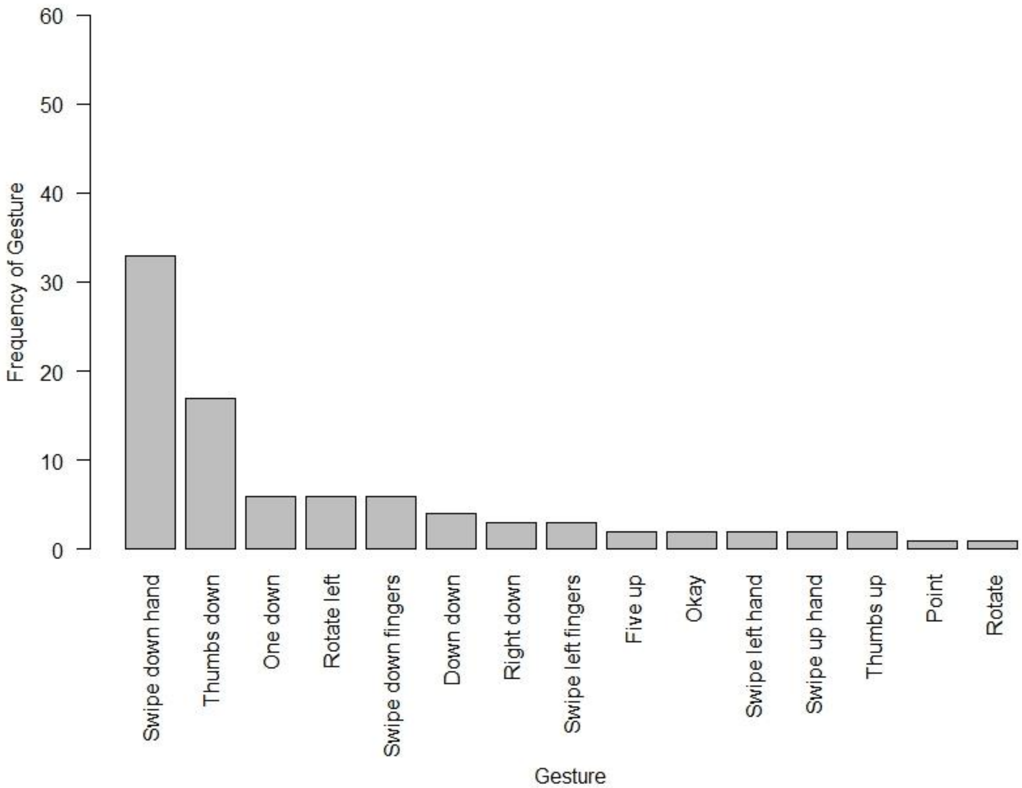


Figure 12. Histogram of gestures chosen for Function 4: Decrease the flow of anesthesia gas

For function 5 (see Figure 4a), "Silence the alarm," the histogram showing gestures chosen is in Figure 13. The most frequently chosen gesture was "swipe left hand" (Figure 8f) and was performed 15 times, and the second most frequently chosen gesture was "swipe right hand" (Figure 8e) and was performed 12 times. The rest of the gestures for function 5 were performed less than 10 times. Although there is little difference between the first and second most frequently chosen gesture in terms of number of times performed, both gestures (swipe left hand and swipe right hand) are a swipe of the hand in the lateral direction.

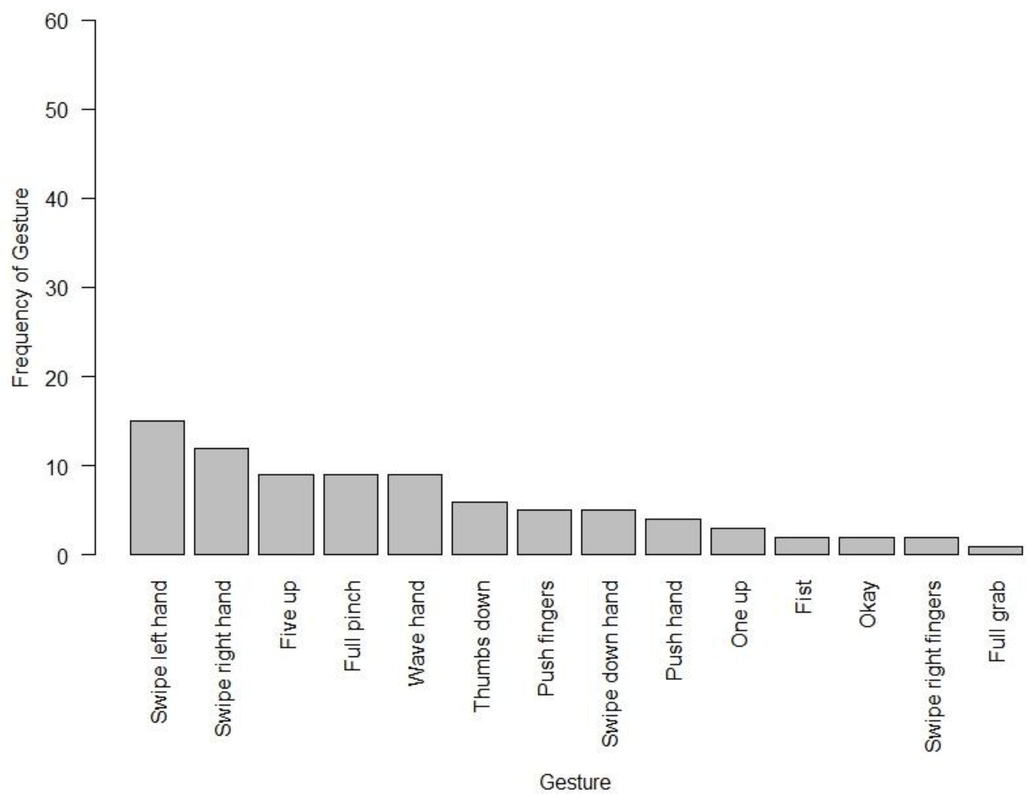


Figure 13. Histogram of gestures chosen for Function 5: Silence the alarm

Function 6 (see Figure 4b) asked to "Acknowledge the message" and the histogram of gestures chosen is in Figure 14. The most frequently chosen gesture was "thumbs up" (Figure 8a) and performed a total of 46 times. The second most frequently chosen gesture was "okay" (Figure 8k) and was performed 16 times. The rest of the functions chosen for function 6 were performed less than 10 times.

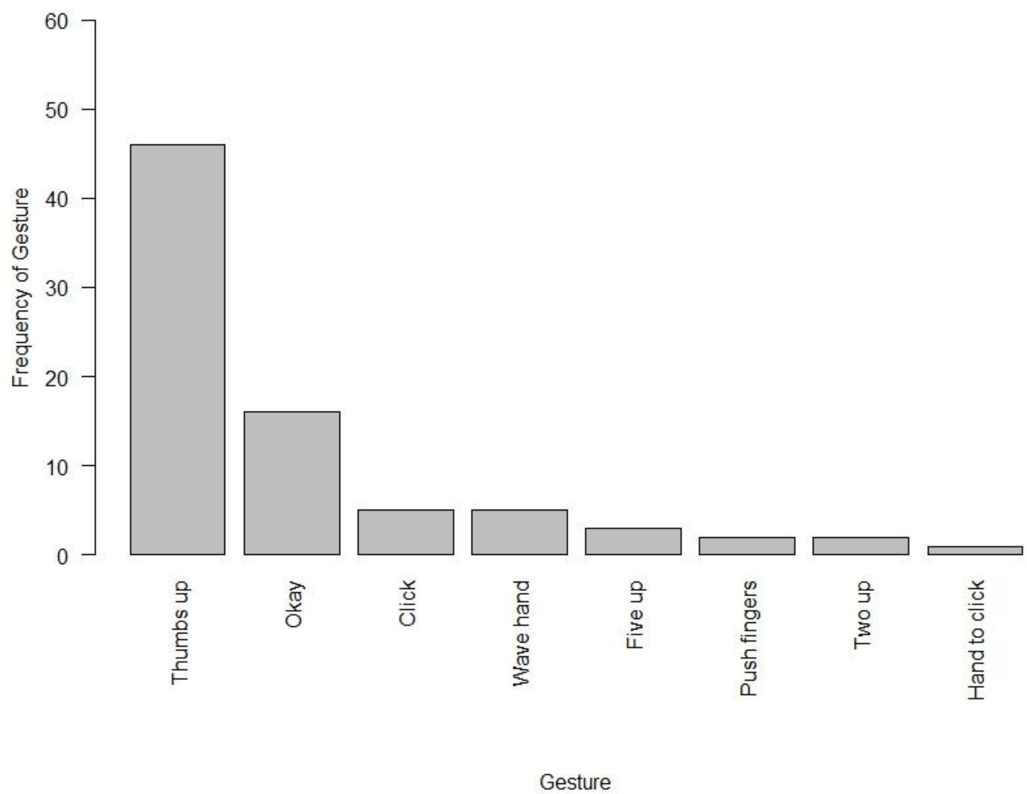


Figure 14. Histogram of gestures chosen for Function 6: Acknowledge the message

Functions 7 (see Figure 5a) and 8 (see Figure 5b) were "Yes/No" type questions asking whether the value of some parameter fell within a range. The histograms of gestures chosen for these functions are shown in Figures 15 and 16, respectively. For Function 7 (Is heart rate normal?), the correct response would be a positive gesture since the heart rate fell within the stated paramters. For function 7, the most frequently chosen gesture was "thumbs up" (Figure 8a) and was performed a total of 54 times. The rest of the chosen gestures were performed less than 10 times. For function 8 (Is pulse oximeter, SpO_2, normal?), the correct response would be a negative response since the pulse oximeter value or the SpO_2 value fell out of the stated paramters. The most frequently chosen gesture was "thumbs up" (Figure 8a) and was performed 25 times, and the second most frequently chosen gesture was "thumbs down" (Figure 8b) and was performed 23 times. The rest of the gestures chosen were performed 10 or less times.

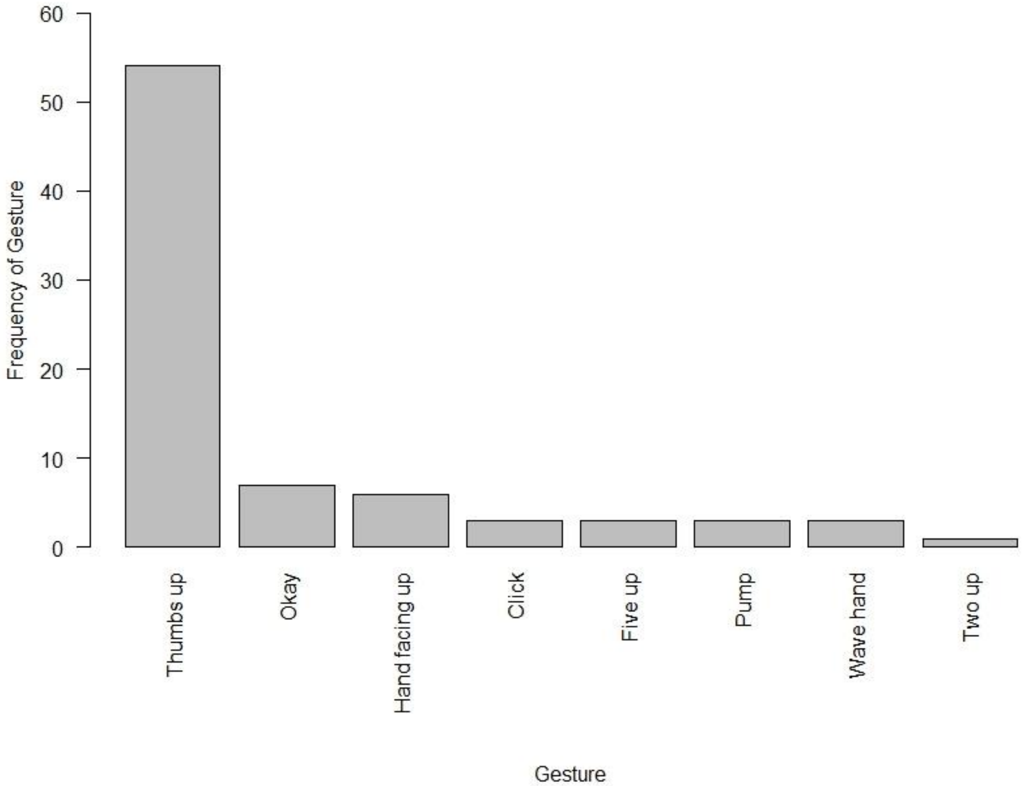


Figure 15. Histogram of gestures chosen for Function 7: Is heart rate normal?

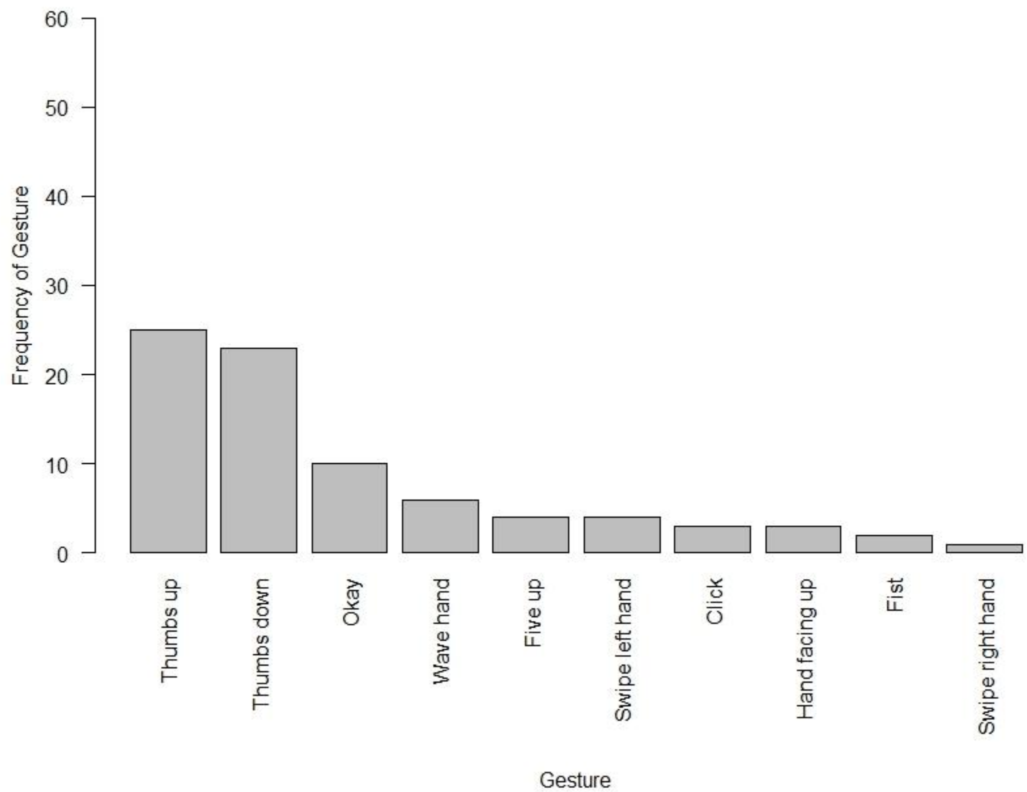


Figure 16. Histogram of gestures chosen for Function 8: Is pulse ox normal?

Function 9 (see Figure 6a) was added as function to be tested as it would be a new functionality of a gestural input system. Function 9 (Select heart rate) was a potential display with four button-like options displayed vertically with heart rate listed as the third option from the top. The histogram of gestures chosen is shown in Figure 17. The most frequently chosen gesture was "push fingers" (Figure 8j) and was performed 25 times. The second most frequently chosen gesture was "three up" (Figure 8l) and was performed 19 times.

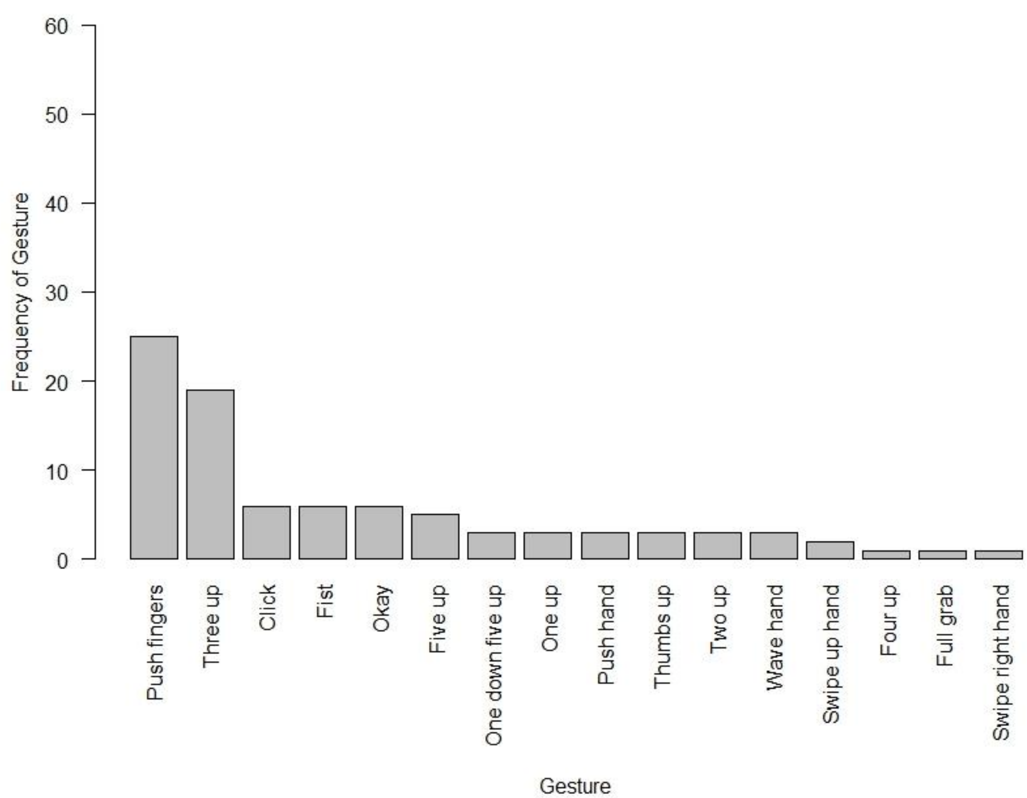


Figure 17. Histogram of gestures chosen for Function 9: Select heart rate (HR).

The last function tested, Function 10 (see Figure 6b), was cancel the request. The histogram for gestures chosen is shown in Figure 18. For this function, the most frequently chosen gesture was “swipe left hand” (Figure 8e) and was performed a total of 25 times. The second most frequently chosen gesture was “thumbs down” (Figure 8b) and was performed 13 times. The third most frequently chosen gesture was “swipe right hand” (Figure 8f) and was performed 12 times. The rest of the gestures performed for function 10 were performed 10 or less times.

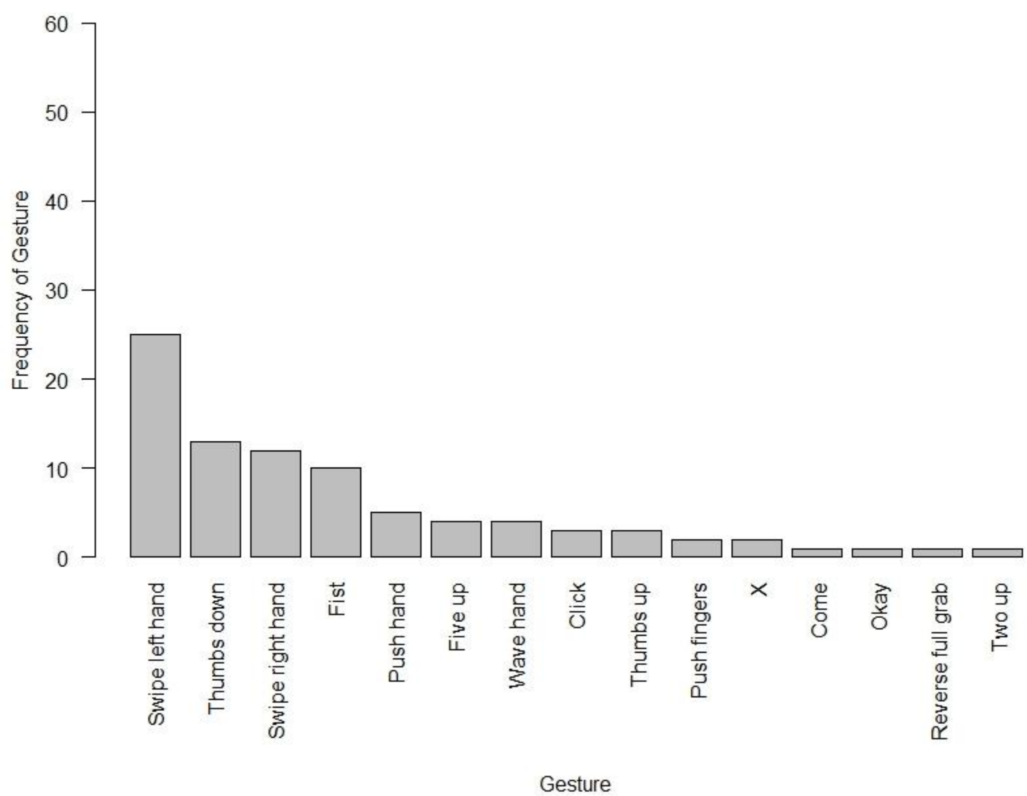


Figure 18. Histogram of gestures chosen for Function 10: Cancel the request

Aim 2 Results

Regression Assumptions

For Aim 1, the only data analyzed was the data from the first block in the experiment to investigate response time at the first instance of a function for a participant. Recall from Chapter 2 that there were several assumptions that needed to be met before moving forward with the mixed model regression analysis, including normality, better performance compared to a fixed model, linearity, independence, homoscedasticity, and no multicollinearity issues. The skew of response times exhibited a positive skew with a value of 2.52, so the data was transformed by taking the natural logarithm of response times. After the transformation, the skew was 0.47. It is recommended that if the skewness of the data is between -0.5 and 0.5, then the distribution of the data can be considered fairly symmetrical (Bulmer, 2012). The ANOVA of the model comparison showed that the linear regression model with "Participant ID" as the random effect performed significantly better than the regression model with only the fixed intercept ($P<0.0001$); therefore, the mixed model was used for the analysis. Since the mixed model was used for analysis, the within-subjects variability is removed as each participant is treated as a random effect; therefore, the assumption of independence of the data is met. The VIF values of this model were calculated and all VIF values were less than 5 indicating that there were no severe multicollinearity issues. Figure 19 shows the model validation plots checking for linearity, homoscedasticity and normality. The residuals plot does not show any curvature indicating that the data follows a linear pattern. The residuals plot also shows homoscedasticity in that the residuals are evenly distributed on either side of the line indicating a residual of 0. The normal Q-Q plot in Figure 5 shows that the data falls on a straight line indicating the data is normally distributed. There were no influential points in the data set as all of the calculated Cook's distances were below the cutoff value. The cutoff value for this dataset was 4/n=4/30=0.133.

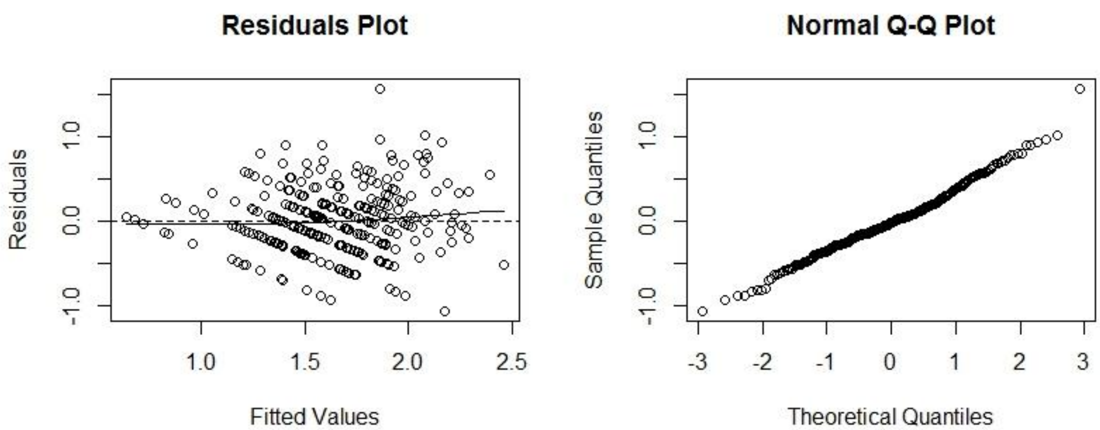


Figure 19. Residuals plot confirms homoscedasticity and linearity and the normal Q-Q plot confirms normality

Cognitive Complexity

The mean response time for functions in Block 1 was 5.90 seconds with a standard deviation of 3.66 seconds. A summary of the mixed effects model is shown in Table 4. The only significant predictors of this model are Function 2, Function 8 and Function 9, and none of the fixed effects were significantly associated with response times.

Table 4. Output summary of mixed effects model.

Variables	Estimate	SD	t-value	P
(Intercept)	0.796	1.157	0.688	0.491
Function 2	-0.238	0.105	-2.259	0.024*
Function 3	-0.153	0.105	-1.455	0.146
Function 4	-0.195	0.105	-1.854	0.064
Function 5	-0.044	0.105	-0.417	0.677
Function 6	-0.054	0.105	-0.514	0.607
Function 7	0.081	0.105	0.770	0.441
Function 8	0.270	0.105	2.572	0.010*
Function 9	0.338	0.105	3.216	0.001*
Function 10	-0.033	0.105	-0.310	0.757
Age	0.035	0.046	0.760	0.447
Gender, Male	0.172	0.203	0.847	0.397
Handedness, Left	-0.113	0.410	-0.276	0.782
Handedness, Right	-0.002	0.362	-0.006	0.995
Education, degree, High School/GED	-0.004	0.240	-0.016	0.987
Education, degree, Masters	-0.069	0.257	-0.269	0.788
Education, Science & Engineering, Yes	0.039	0.156	0.245	0.807
Video Game Use, Yes	0.020	0.190	0.108	0.914
Virtual Reality Gaming Experience, Yes	-0.057	0.151	-0.378	0.705

Note: * indicates that the variable was significantly associated with response time ($P<0.05$)

Recall from Chapter 2 that the mixed effects model only reveals which functions are significantly associated with response times in comparison to the reference function, which in this case it is Function 1. The Tukey contrasts of all pairwise comparisons of functions is summarized in Table 5; only the coefficients of the contrasts with significant associations are included in the table. The table can be read from row to column. For example, Function 9 takes significantly longer compared to Function 1. The diagonal represents comparisons against the same function which is not feasible. The pairwise comparisons to the left of the diagonal represent all pairwise compairons. The comparisons to the right of the diagonal were not completed because they would echo the results from the left side, where the only difference is a flipped sign. For example, Function 9 takes significantly longer compared to Function 10. Only the parwise comparisons which show a significant effect are included in the table (P<0.05).

Table 5. Coefficients of Tukey Contrasts for all function pairwise comparisons in regression model. (-) indicates an insignificant association, filled in cells indicate impossible comparison, and blank cells indicate repetitive data that is not reported.

Function	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1										
2	-									
3	-	-								
4	-	-	-							
5	-	-	-	-						
6	-	-	-	-	-					
7	-	-	-	-	-	-				
8	-	0.53	0.44	0.48	-	-	-			
9	0.36	0.60	0.51	0.55	0.41	0.41	-	-		
10	-	-	-	-	-	-	-	-	-0.39	

Function 9 takes significantly longer to choose a gesture compared to functions 1, 2, 3, 4, 5, 6, and 10 suggesting that the cognitive load is respectively higher for Function 9 compared to these functions. Function 8 takes significantly longer than functions 2, 3, and 4 also indicating that the cognitive load is respectively higher for Function 8 compared to functions 2, 3, and 4.

General Findings

A compilation of the results from Aim 1 and Aim 2 are summarized in Table 6. The ideal situation would be a function to pose one clear intuitive gesture without any cognitive complexity concerns, and six functions demonstrated intuitive gesture mappings with no cognitive complexity issues as shown in the first six rows of Table 6. Two functions demonstrated no intuitive gesture mappings but some higher level gestural mapping without any cognitive complexity issues as shown in the next couple rows of Table 6. The last two functions demonstrated no intuitive mappings, no commonalities and multiple cognitive complexity issues as shown in the last two rows of Table 6.

Table 6. Summary of Aim 1 and Aim 2 results- Functions with intuitive gesture mapping, higher level mappings if there is no intuitive gesture, and functions that pose a cognitive complexity concerns.

Function	Intuitive Gesture Mapped	If No Intuition, Higher Level Mapping?	Cognitive Complexity Concerns?
Start the flow	Thumbs up	-	-
Inc. the flow	Swipe up hand	-	-
Dec. the flow	Swipe down hand	-	-
Ack. the message	Thumbs up	-	-
Heart rate normal?	Thumbs up	-	-
Cancel the message	Swipe left hand	-	-
Stop the flow	Five up/ Push hand	Five up posture	-
Silence alarm	Swipe left hand/ Swipe right hand	Lateral hand movement	-
Pulse ox normal?	Thumbs up/ thumbs down	-	Yes
Select HR.	Push fingers/ three up	-	Yes

CHAPTER 4. DISCUSSION

Gestures are a natural means of human-human communication, but gestures are also potentially a means of human-computer communication (Karam & Schraefel, 2005). The objective of this research was to map 3D gestural inputs to traditional touchscreen interface designs within the context of anesthesiology. The two aims of this research sought to identify intuitive gesture-function mappings and identify any gesture-function mappings that demonstrated cognitive complexity issues. Previous research has shown that users prefer gestural input systems to incorporate gestures that are elicited from end users (Morris et al., 2010); therefore, this work adopted and modified a user elicitation approach (Nielsen et al., 2004) to fulfill the objective and aims of this research. The results of this work indicate that gestural input technology is fit for application for anesthesiology in the OR. There was an intuitive mapping for functions 1, 3, 4, 6, 7, and 10. All of these functions demonstrated a specific gesture that was performed more frequently than the other chosen gestures for that respective function. It is important to note that there was also not a significant difference in response times for all pairwise comparisons within this group of functions. These functions demonstrated the ideal situation where there is an intuitive mapping and no cognitive complexity concerns.

As for the functions that did not fit the above mentioned ideal scenario, there were not any major concerns associated with the rest of the functions that merited unfitness of gestural input technology. Functions 2 and 5 are unique in that there were not any significant differences in response times in all pairwise comparisons, but there was not a clear gesture that stuck out as being performed more than any other gesture. As stated in the Aim 1 Results section in Chapter 3, the top two gestures for Function 2 were “five up” (Figure 8c) and “push hand” (Figure 8d), so although there was not a most frequent gesture chosen, there is some commonality in these two gestures in that both incorporate the “five up” posture, as shown in Table 6. If these two would be combined, it would result in 31 total gestures and the new second top gesture (“fist”) was only performed 12 times. Function 5 also did not exhibit a clear top gesture, but there is also some commonality between the two top chosen gestures, “swipe right hand” (Figure 8f) and “swipe left hand” (Figure 8e), in that both of these gestures demonstrate a lateral hand movement as shown in Table 6. If these two were to be combined, it would result in a total of 27 gestures, and the new second top gesture (“five up”) was only performed 9 times. The results of functions 2 and 5 do not necessarily demonstrate that gestural input technology is unfit for these functions, but the results rather suggest that the gestural classification scheme used in the analysis did not account for higher level commonalities of gestures. There is still potential for gestural input technology for use in these functions, but the design of the gesture-function mapping would need to incorporate the higher level gesture type rather than a lower level, specific gesture.

Function 8 was the function that asked the participant to evaluate the pulse oximeter and the correct answer should have been a negative response. The top two gestures chosen for this function respectively were “thumbs up” (Figure 8a) and “thumbs down” (Figure 8b), which are opposite gestures. Function 8 also demonstrated longer response times, and subsequently had higher cognive load, compared to three functions (2, 3, and 4). This may be because the participants were not anesthesia providers and may be less familiar with the pulse oximeter by choosing “thumbs up” (Figure 8a). Because of this observation, this may explain why some participants incorrectly evaluated the pulse oximeter parameter. Like functions 2 and 5, Function 8 is also not of concern to gestural input technology with the current results because these results may have been observed due to the participants having a lack of clinical knowledge. However,it cannot be assumed that the participants who chose “thumbs up” chose an incorrect gesture due to unfamiliarity of the pulse oximeter. The “thumbs up” responses could be because the participant was responding to the fact that they saw the message and not actually evaluating the parameter. For future studies, it would be valuable to add a think-aloud protocol to determine how participants approach the specific function. In order to test Function 8 again to determine fitness for gestural input technology, it would need to be tested with actual anesthesia providers and incorporate a think-aloud protocol.

Function 9’s performance was unlike any of the other functions in that the response times were longer than 7 of the other 9 functions, does not demonstrate any intuitive mappings, and there was not any commonalities between the top gestures as these two gestures create two completely different interactions. One of the top two gestures was “push fingers” (Figure 8j), and this gesture is an interaction where the participant is reaching towards the computer screen where heart rate is fixated in space as if trying to push that button, similar to a deictic gesture. The other top gesture chosen was “three up” (Figure 8l), and this gesture interaction is more symbolic as heart rate was the third button down from the top. This function did not have a most frequent intuitive gesture, but rather there are two completely different gestures that emerged as being a natural interaction for this group of users. This may be because Function 9 would be a new functionality of a gestural input system and is therefore more unfamiliar to the participants than some of the other functions, These results also do not pose any major concerns for fitness of gestural input technology as it is a new functionality that gestural input technology would incorporate. Future research and further development of the technology and gestural interaction will address whether there could be multiple intuitive gestures.

As for the gestures that did not demonstrate an intuitive gesture mapping, this motivates the notion to classify gestures through a different approach, one that incorporates different properties of gestures to identify commonalities at different levels. This motivation is especially seen in the two gestures that demonstrated some commonality between gestures chosen despite not having an intuitive gesture mapping. The gesture dictionary used in the study lacks the separation between gesture and context. For example, the gesture "okay" (Figure 8k) could have been performed by participants who were meaning to do a gesture indicating "three." If there was a gesture taxonomy that was based on the biomechanics of gestures, this would support finding different levels of commonalities between the gestures chosen for a function. Overall, a majority of the functions demonstrated intuitive mappings without cognitive complexity issues and the other functions did not pose any major concerns regarding gestural input technology fitness which gives reason to believe gestural input technology has a place for anesthesiology in the OR.

It is noteworthy to mention that three of the functions demonstrated the same intuitive gesture mapping. Functions 1 (Start the flow of anesthesia gas), 6 (Acknowledge the message) and 7 (Is heart rate normal?) were all mapped to the "thumbs up" (Figure 8a) gesture. In a study done by Pereira et al. (2015) there were also gestures that were mapped to more than one function. If these mappings hold true throughout development of gestural systems, context sensitivity becomes much more important as the technology needs to be capable of understanding the context around which function a gesture is appropriate for. However, unlike the work from Pereira et al. (2015), this study was done within a particular context and was not general human-computer interaction tasks.

Only two functions posed concerns with cognitive complexity. Function 8 was potentially complex for the study participants due to a lack of clinical knowledge, and Function 9 was potentially complex due to it being a new functionality of gestural input technology. It is also interesting that none of the other fixed effects in the regression model were associated with response time indicating that gestural input technology can be used for a variety of users.

Overall this work has shown that gestural input technology could be applied to anesthesiology in the OR. The market demand for this technology was the infection control concern within anesthesiology in the OR. HAIs create an unnecessary burden on the healthcare system (Klevens et al., 2007; U.S. Department of Health and Human Services, 2013), and anesthesia provider's hands have been shown to play a key role in bacterial transfer (Loftus et al., 2012). By incorporating gestural input technology, anesthesia providers can reduce the number of surfaces, equipment and other devices they come in contact with, which leads to potentially reducing the risk of infection. Although the results of this study favor gestural input technology, it cannot be concluded that gestural input technology will generate a positive impact with respect to HAIs. Future research needs to build off of the current work to ultimately produce an impact that reduces the burden of HAIs.

Limitations

There are several limitations associated with this study. Although the experimental setting attempts to replicate aspects of an anesthesia work environment, the participants' responses does not reflect the behavior of anesthesia providers since none of the participants work within anesthesiology. Although this is so, it actually benefits this research to not have anesthesiologists during these beginning stages of design and development as the students who participated in the research provided a true bottom-up approach as none of the students were familiar with the application and do not have any prior knowledge of how the system should work. Another limitation with this study which may have affected the gestures chosen for Function 7 ("Is heart rate normal?") and Function 8 ("Is pulse ox normal?") is the fact heart rate is listed in green on the function slide and pulse ox is listed in yellow on the function slide. Traditionally, the color green has been considered to embody a more positive meaning than the color yellow (Madden, Hewett, & Roth, 2000), so the participants could have demonstrated a tendency to choose a more positive response for the green value (i.e., heart rate). Additionally, the Wizard of

Oz techniques used during the experiment may impact the results as the technology was simulated via the experimenter and not by the technology, but the Wizard of Oz methodology allowed for a quicker advancement to the next function via the wizard than if the technology recognized the gesture on its own. This methodology also eliminated any potential user frustration with the reliability of the technology. The practice session of the study is also a limitation as it may influence the gestures chosen during the experiment, but this was not entirely shown in the results as there were 42 unique gestures performed and the participants only practiced 14 gestures. Also, no data with respect to the participant's culture was collected in the experiment. Some gestures have been described to have a cultural dependency (Efron, 1941), but this experiment did not account for cultural differences between meanings of gestures. Lastly this study is exploratory and seeks whether gestural input technology has a place in the OR for anesthesiology so different results may be generated in a larger study.

Future Research

Supported by the results presented in Chapter 3, future research should investigate a biomechanical approach to classifying gestures for gestural input systems in order to identify the highest level of commonality or level of agreement in gestures chosen. For example, it may be beneficial to understand if the highest level of agreement between users is in the number of fingers used, static or dynamic movement, posture chosen, or orientation of the palm. Doing so may point towards more valuable results in terms of how to design a gesture vocabulary set to use for anesthesiology in the OR.

As the intuitive gesture-function mappings are determined through the biomechanical classification approach and testing with anesthesia providers, the logical next step is to test the physical limitations of the gestures in terms of musculoskeletal fatigue. If an intuitive gesture mapping turns out to be physically difficult to perform over a long period of time, then the design of the gestural input system needs to be restructured to account for any musculoskeletal concerns. In addition to the physical limitations, the limitations of the technology also need to be tested. For example, once there is a final set of gesture-function mappings, studies will be completed to test the usability of the gestural input system in terms of workload, reliability of the technology, ease of use, and usefulness of the system.

Future research should also replicate this study with anesthesia providers. Students were used in these beginning stages of research to determine whether or not gestural input technology has a place within this context, but as it has been determined that gestural input technology does show promising results among students, this study should be replicated with anesthesia providers to determine if the same results hold true.

Impacts and Implications

With the release of new technologies such as vision-based, 3-D gestural input technology, the way humans interact with technology is dramatically changed. It is important to continue with the future research in developing this technology to maximize performance of the system, as well as to understand the human factors implications for the use of this technology. To continually ensure success of a gestural input system, it is important to develop these gestural input systems within a certain contextual perspective

(Ardito et al., 2014). This research is unique as a gestural input system for anesthesiology has not been investigated. The results of this research will contribute to the application of gestural input systems in the anesthesia work environment in the OR. Anesthesiology has an infection control concern in the OR, and gestural input technology can help minimize the risk of infection to the patient by creating a touchless environment. Anesthesia providers may reduce the number of surfaces they come in contact with by using gestures to perform certain functions when interacting with computers and other devices. Gestural input systems are capable of improving patient safety because minimizing touching will ultimately break the pattern of bacterial transmission from the patient to the anesthesia environment and protect sterility in the OR.

REFERENCES

Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., … Koh, J. (2012). Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. *Microsoft Research TechReport MSR-TR-2012-111*.

Anderson, F., & Bischof, W. F. (2013). Learning and performance with gesture guides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1109–1118). ACM.

Ardito, C., Costabile, M. F., & Jetter, H.-C. (2014). Gestures that people can understand and use. *Journal of Visual Languages & Computing*, *25*(5), 572–576. http://doi.org/http://dx.doi.org/10.1016/j.jvlc.2014.07.002

Barsuk, J. H., Cohen, E. R., Feinglass, J., McGaghie, W. C., & Wayne, D. B. (2009). Use of simulation-based education to reduce catheter-related bloodstream infections. *Archives of Internal Medicine*, *169*(15), 1420–1423.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1).

Baudel, T., & Beaudouin-Lafon, M. (1993). Charade: remote control of objects using free-hand gestures. *Communications of the ACM*, *36*(7), 28–35.

Biddle, C., & Shah, J. (2012). Quantification of anesthesia providers' hand hygiene in a busy metropolitan operating room: what would Semmelweis think? *American Journal of Infection Control*, *40*(8), 756–759.

Birnbach, D. J., Rosen, L. F., Fitzpatrick, M., Carling, P., Arheart, K. L., & Munoz-Price, L. S. (2015). Double Gloves: A Randomized Trial to Evaluate a Simple Strategy to Reduce Contamination in the Operating Room. *Anesthesia & Analgesia*, *120*(4). Retrieved from http://journals.lww.com/anesthesia-analgesia/Fulltext/2015/04000/Double_Gloves___A_Randomized_Trial_to_Evaluate_a.23.aspx

Birnbach, D. J., Rosen, L. F., Fitzpatrick, M., Carling, P., & Munoz-Price, L. S. (2015). The Use of a Novel Technology to Study Dynamics of Pathogen Transmission in the Operating Room. *Anesthesia & Analgesia*, *120*(4). Retrieved from http://journals.lww.com/anesthesia-analgesia/Fulltext/2015/04000/The_Use_of_a_Novel_Technology_to_Study_Dynamics_of.22.aspx

Bizzotto, N., Costanzo, A., Bizzoto, L., Regis, D., Sandri, A., & Magnan, B. (2014). Leap motion gesture control with OsiriX in the operating room to control imaging: first experiences during live surgery. *Surgical Innovation*, *1*.

Bulmer, M. G. (2012). *Principles of statistics*. BOOK, Courier Corporation.

Clark, C., Taenzer, A., Charette, K., & Whitty, M. (2014). Decreasing contamination of the anesthesia environment. *American Journal of Infection Control*, *42*(11), 1223–1225. http://doi.org/http://dx.doi.org/10.1016/j.ajic.2014.07.016

Comer, A., Harris, A. D., Shardell, M., Braun, B., Belton, B. M., Wolfsthal, S. D., … Sulis, C. (2011). Web-based training improves knowledge about central line bloodstream infections. *Infection Control & Hospital Epidemiology*, *32*(12), 1219–1222.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, *19*(1), 15–18.

Dong, H., Danesh, A., Figueroa, N., & El Saddik, A. (2015). An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access*, *3*, 543–555.

Efron, D. (1941). *Gesture, race and culture: a tentative study of the spatio-temporal and" linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions* (Vol. 9). Mouton.

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. *Studies in Dyadic Communication*, 153–175.

Freeman, D., Benko, H., Morris, M. R., & Wigdor, D. (2009). ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (pp. 165–172). ACM.

Horsky, J., Kaufman, D. R., Oppenheim, M. I., & Patel, V. L. (2003). A framework for analyzing the cognitive complexity of computer-assisted clinical ordering. *Journal of Biomedical Informatics*, *36*(1), 4–22. http://doi.org/10.1016/S1532-0464(03)00062-5

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.

Höysniemi, J., Hämäläinen, P., & Turkki, L. (2004). Wizard of Oz Prototyping of Computer Vision Based Action Games for Children. In *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community* (pp. 27–34). New York, NY, USA: ACM. http://doi.org/10.1145/1017833.1017837

Hurka, F., Wenger, T., Heininger, S., & Lueth, T. C. (2011). Method, accuracy and limitation of computer interaction in the operating room by a navigated surgical instrument. In *2011 Annual International Conference of the IEEE Engineering in*

Medicine and Biology Society (pp. 2144–2147). IEEE.

Jacob, M. G., & Wachs, J. P. (2014). Context-based hand gesture recognition for the operating room. *Pattern Recognition Letters*, *36*, 196–203.

Jacob, M. G., Wachs, J. P., & Packer, R. A. (2013). Hand-gesture-based sterile interface for the operating room using contextual cues for the navigation of radiological images. *Journal of the American Medical Informatics Association*, *20*(1), 183–186.

Karam, M., & Schraefel, M. C. (2005). A Taxonomy of Gestures in Human Computer Interactions. University of Southampton. *Computing Service*.

Kendon, A. (1988). How gestures can become like words. *Cross-Cultural Perspectives in Nonverbal Communication*, *1*, 131–141.

Klevens, R. M., Edwards, J. R., Richards, C. L., Horan, T. C., Gaynes, R. P., Pollock, D. A., & Cardo, D. M. (2007). Estimating Health Care-Associated Infections and Deaths in U.S. Hospitals, 2002. *Public Health Reports (1974-)*, *122*(2), 160–166. Retrieved from http://www.jstor.org/stable/20057111

Koff, M. D., Brown, J. R., Marshall, E. J., O'Malley, A. J., Jensen, J. T., Heard, S. O., … Loftus, R. W. (2016). Frequency of Hand Decontamination of Intraoperative Providers and Reduction of Postoperative Healthcare-Associated Infections: A Randomized Clinical Trial of a Novel Hand Hygiene System. *Infection Control and Hospital Epidemiology*, 1–8.

Koff, M. D., Loftus, R. W., Burchman, C. C., Schwartzman, J. D., Read, M. E., Henry, E. S., & Beach, M. L. (2009). Reduction in intraoperative bacterial contamination of peripheral intravenous tubing through the use of a novel device. *The Journal of the American Society of Anesthesiologists*, *110*(5), 978–985.

Loftus, R. W., Brown, J. R., Koff, M. D., Reddy, S., Heard, S. O., Patel, H. M., … Jensen, J. T. (2012). Multiple reservoirs contribute to intraoperative bacterial transmission. *Anesthesia & Analgesia*, *114*(6), 1236–1248.

Loftus, R. W., Muffly, M. K., Brown, J. R., Beach, M. L., Koff, M. D., Corwin, H. L., … Yeager, M. P. (2011). Hand contamination of anesthesia providers is an important risk factor for intraoperative bacterial transmission. *Anesthesia & Analgesia*, *112*(1), 98–105.

Madden, T. J., Hewett, K., & Roth, M. S. (2000). Managing Images in Different Cultures: A Cross-National Study of Color Meanings and Preferences. *Journal of International Marketing*, *8*(4), 90–107. JOUR. Retrieved from http://www.jstor.org.libproxy.clemson.edu/stable/25048831

Marschall, J., Mermel, L. A., Fakih, M., Hadaway, L., Kallen, A., O'Grady, N. P., … Maragakis, L. L. (2014). Strategies to prevent central line-associated bloodstream infections in acute care hospitals: 2014 update. *Infection Control & Hospital Epidemiology*, *35*(S2), S89–S107.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Mewes, A., Saalfeld, P., Riabikin, O., Skalej, M., & Hansen, C. (2016). A gesture-controlled projection display for CT-guided interventions. *International Journal of Computer Assisted Radiology and Surgery*, *11*(1), 157–164.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (n.d.). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2015. *R Package Version*, 1–6.

Miller, R. D., & Pardo, M. (2011). *Basics of anesthesia* (6th ed.). Philadelphia, PA: Elsevier. http://doi.org/10.1017/CBO9781107415324.004

Morris, M. R., Wobbrock, J. O., & Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010* (pp. 261–268). CONF, Canadian Information Processing Society.

Munoz-Price, L. S., Lubarsky, D. A., Arheart, K. L., Prado, G., Cleary, T., Fajardo-Aquino, Y., … Birnbach, D. J. (2013). Interactions between anesthesiologists and the environment while providing anesthesia care in the operating room. *American Journal of Infection Control*, *41*(10), 922–924. http://doi.org/http://dx.doi.org/10.1016/j.ajic.2013.03.306

Nielsen, M., Störring, M., Moeslund, T., & Granum, E. (2004). A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In A. Camurri & G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction SE - 38* (Vol. 2915, pp. 409–420). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-24598-8_38

Nieuwenhuis, R., te Grotenhuis, H. F., & Pelzer, B. J. (2012). Influence. ME: tools for detecting influential data in mixed effects models.

O'Hara, K., Gonzalez, G., Sellen, A., Penney, G., Varnavas, A., Mentis, H., … Dastur, N. (2014). Touchless interaction in surgery. *Communications of the ACM*, *57*(1), 70–77.

Pereira, A., Wachs, J. P., Park, K., & Rempel, D. (2015). A user-developed 3-D hand gesture set for human–computer interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(4), 607–621.

Pickering, C. A., Burnham, K. J., & Richardson, M. J. (2007). A research study of hand gesture recognition technologies and applications for human vehicle interaction. In *3rd Conf. on Automotive Electronics*. Citeseer.

Pittet, D., Hugonnet, S., Harbarth, S., Mourouga, P., Sauvan, V., Touveneau, S., & Perneger, T. V. (2000). Effectiveness of a hospital-wide programme to improve compliance with hand hygiene. *The Lancet*, *356*(9238), 1307–1312. http://doi.org/http://dx.doi.org/10.1016/S0140-6736(00)02814-2

Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, *43*(1), 1–54. http://doi.org/10.1007/s10462-012-9356-9

Rowlands, J., Yeager, M. P., Beach, M., Patel, H. M., Huysman, B. C., & Loftus, R. W. (2014). Video observation to map hand contact and bacterial transmission in operating rooms. *American Journal of Infection Control*, *42*(7), 698–701.

Sax, H., Allegranzi, B., Uçkay, I., Larson, E., Boyce, J., & Pittet, D. (2007). "My five moments for hand hygiene": a user-centred design approach to understand, train, monitor and report hand hygiene. *Journal of Hospital Infection*, *67*(1), 9–21. http://doi.org/http://dx.doi.org/10.1016/j.jhin.2007.06.004

Schroder, S., Loftfield, N., Langmann, B., Frank, K., & Reithmeier, E. (2014). Contactless operating table control based on 3D image processing. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. http://doi.org/10.1109/EMBC.2014.6943610

Stackhouse, R. A., Beers, R., Brown, D., Brown, M., Greene, E., McCann, M. E., & Tompkins, B. M. (2011). *Recommendations for Infection Control for the Practice of Anesthesiology*. *American Society of Anesthesiologists* (3rd ed.). Park Ridge, IL. Retrieved from https://asahq.org/~/media/sites/asahq/files/public/resources/asa committees/recommendations-for-infection-control-for-the-practice-of-anesthesiology-(1).pdf?la=en

Stern, H. I., Wachs, J. P., & Edan, Y. (2006). Human Factors for Design of Hand Gesture Human - Machine Interaction. *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on*. http://doi.org/10.1109/ICSMC.2006.384767

The Research Committee of the Society of Healthcare Epidemiology of America. (2010). Enhancing Patient Safety by Reducing Healthcare-Associated Infections: The Role of Discovery and Dissemination. *Infection Control and Hospital Epidemiology : The Official Journal of the Society of Hospital Epidemiologists of America*, *31*(2), 118–123. http://doi.org/10.1086/650198

U.S. Department of Health and Human Services. (2013). National Action Plan to prevent health care-associated infections: road map to elimination. *Washington, DC: USDHHS*. Retrieved from http://www.health.gov/hai/pdfs/hai-action-plan-executive-summary.pdf

Ulrich, K. T., & Eppinger, S. D. (2012). *Product Design and Development* (5th ed.). New

York, NY: McGraw-Hill.

Van der Meer, T., Te Grotenhuis, M., & Pelzer, B. (2010). Influential cases in multilevel modeling a methodological comment. *American Sociological Review*, *75*(1), 173–178.

Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, *54*(2), 60–71.

Wachs, J. P., Stern, H. I., Edan, Y., Gillam, M., Handler, J., Feied, C., & Smith, M. (2008). A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association*, *15*(3), 321–323.

Weber, D. J., Anderson, D., & Rutala, W. A. (2013). The role of the surface environment in healthcare-associated infections. *Current Opinion in Infectious Diseases*, *26*(4), 338–344.

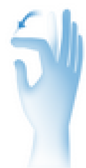
Wigdor, D., & Wixon, D. (2011). *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier.

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1083–1092). CONF, ACM.

World Health Organization. (2009). *WHO guidelines on hand hygiene in health care: first global patient safety challenge. Clean care is safer care*. World Health Organization.

APPENDICES

Appendix A. Gestures in Intel RealSense Software Development Kit

	Gesture	Description
1		Open hand facing the camera, moves the index finger quickly toward the palm center
2		All fingers folded into a fist. The fist can be in different orientations as long as the palm is in the general direction of the camera.
3		All fingers extended and touching the thumb. The pinched fingers can be anywhere between pointing directly to the screen or in profile.
4		Hand open, facing the camera.
5		Hand with palm facing the camera, moves down and immediately back to the starting position.

6		Hand with palm facing the camera, moves left and immediately back to the starting position.
7		Hand with palm facing the camera, moves right and immediately back to the starting position
8		Hand with palm facing the camera, moves up and immediately back to the starting position.
9		A hand in a natural relaxed pose is moved forward as if pressing button.
10		Hand closed with thumb pointing down.
11		Hand closed with thumb pointing up.

12		Hand open with thumb and index finger touching each other.
13		Hand closed with index finger and middle finger pointing up.
14		An open hand facing the screen. This gesture can include any number of repetitions.

Appendix B. Gesture dictionary used for classification of gestures

Static Gestures

Gesture Name	*Description*
Okay	static okay sign – pinky, ring, and middle fingers pointed up, pointer finger is touching thumb
Fist	all fingers are together forming a static fist
One up	Static position of the pointer finger upwards and other fingers tucked into palm
Two up	static position of the pointer and middle finger pointed upwards and other fingers tucked into the palm
Three up	static position of the pointer, middle, and ring fingers pointed upwards and the pinky and thumb tucked into the palm
Four up	static position of the pointer, middle, ring, and pinky fingers pointed upwards and thumb tucked into the palm
Five up	all fingers spread out in a static position
One down	Static position of the pointer finger down and other fingers tucked into palm
Two down	static position of the pointer and middle finger pointed down and other fingers tucked into the palm
Three down	static position of the pointer, middle, and ring fingers pointed down and the pinky and thumb tucked into the palm
Four down	static position of the pointer, middle, ring, and pinky fingers pointed down and thumb tucked into the palm
Five down	static position of all fingers pointed down
One right	Static position of the pointer finger pointed right and other fingers tucked into palm
Two right	static position of the pointer and middle finger pointed right and other fingers tucked into the palm
Three right	static position of the pointer, middle, and ring fingers pointed right and the pinky and thumb tucked into the palm
Four right	static position of the pointer, middle, ring, and pinky fingers pointed right and thumb tucked into the palm
Five right	static position of all fingers pointed right
One left	Static position of the pointer finger left and other fingers tucked into palm
Two left	static position of the pointer and middle finger pointed left and other fingers tucked into the palm
Three left	static position of the pointer, middle, and ring fingers pointed left and the pinky and thumb tucked into the palm
Four left	static position of the pointer, middle, ring, and pinky fingers pointed left and thumb tucked into the palm
Five left	static position of all fingers pointed left

Hand facing up	Static position of all fingers spread out and palm facing up
Hand facing down	Static position of all fingers spread out and palm facing down
Hand facing left	Static position of all fingers spread out and palm facing left
Hand facing right	Static position of all fingers spread out and palm facing right
Thumbs down	a static position of thumb pointed down and all other fingers tucked
Thumbs Up	a static position of thumbs pointed up and all other fingers tucked

Dynamic Gestures

Gesture Name	*Description*
Click	Pointer and thumb start in open position and then move to a closed position where the tips of fingers touch
Five up to click	"five up" gesture moves to okay gesture
Full pinch	all fingers start spread out and dynamically come together to form a pinching of all fingers
Full grab	all fingers start spread out in "five up" gesture and dynamically come together to form the fist gesture
Reverse full grab	All fingers start in fist and dynamically spread out to "five up" gesture
Push hand	dynamic movement of the "palm" gesture – there is movement towards the camera and back towards the body. In this movement, there is not a clear static position of the palm
Push fingers	dynamic movement of the fingers – there is movement towards the camera and back towards the body. In this movement, there is not a clear static position of the fingers
Swipe up hand	a movement of the hand (spread fingers) in the upward direction at any orientation of the hands and fingers
Swipe up fingers	a movement of the hands and fingers (in any position but spread fingers) in the upward direction at any orientation of the hands and fingers
Swipe down hand	a movement of the hands and fingers in the downward direction at any orientation of the hands and fingers
Swipe down fingers	a movement of the hands and fingers (in any position but spread fingers) in the upward direction at any orientation of the hands and fingers
Swipe left hand	a movement of the hands and fingers in the leftward direction at any orientation of the hands and fingers
Swipe left fingers	a movement of the hands and fingers (in any position but spread fingers) in the upward direction at any orientation of the hands and fingers

Swipe right hand	a movement of the hands and fingers in the rightward direction at any orientation of the hands and fingers
Swipe right fingers	a movement of the hands and fingers (in any position but spread fingers) in the upward direction at any orientation of the hands and fingers
Down Down	a double movement downwards – for example, thumbs down twice
Up up	a double movement upwards – for example, thumbs up twice
Left left	a double movement left – for example, thumbs up twice
Right right	a double movement right – for example, thumbs up twice
Up right	a movement from orientation up to the right – for example, thumbs up starts up and then goes right
Up left	a movement from orientation up to the left – for example, thumbs up starts up and then goes left
Right down	a movement from orientation right to down – for example, thumbs up starts right and then goes down
Right up	a movement from orientation right to up – for example, thumbs up starts right and then goes up
Left down	a movement from orientation left to down – for example, thumbs up starts left and then goes down
Left up	a movement from orientation left to up – for example, thumbs up starts left and then goes up
Down right	a movement from orientation left to up – for example, thumbs up starts down and then goes right
Down left	a movement from orientation left to up – for example, thumbs up starts down and then goes left
Wave hand	a combination of swipe left palm and swipe right palm. Can be any number of swipe left palm, swipe right palm combinations
Wave finger	a combination of swipe left finger and swipe right finger. Can be any number of swipe left finger, swipe right finger combinations
Pump	Multiple full pinch movements
rotate	The hands and fingers make a rotation motion back and forth
Rotate right	the hands and fingers make a rotation to the right (clockwise)
Rotate left	the hands and fingers make a rotation to the left (counterclockwise)
X	Diagonal hand movements to form an X
One down five up	A movement of the fingers down and then the “five up” gesture
One up five up	A movement of the fingers up and then the “five up” gesture
One left five up	A movement of the fingers left and then the “five up” gesture
One right five up	A movement of the fingers right and then the “five up” gesture
Come	Hand starts with palm facing up and there is a movement towards the body of the pointer, middle, ring, and pinky finger