Improving Robotic Surgery Training with Bimanual Wrist Squeezing Haptic Feedback

Zachary Patterson[1], Brett Wolfinger[2], Katherine J. Kuchenbecker[3], and Jeremy D. Brown[4]

***Abstract*— This paper details an experiment performed to evaluate the potential of haptic feedback to improve learning of a robotic surgery task. As Robotic Minimally Invasive Surgery (RMIS) becomes more common, it becomes more important to properly train surgeons to use surgical robots such as the da Vinci. Because standard robotic surgery platforms do not provide touch cues, learning to use them is challenging. If haptic (tactile) feedback is provided as a training tool, we predict that trainees will learn more quickly by associating the haptic feedback with what they see. To observe this, we conducted an experiment with $N = 23$ novice participants from the general population of Johns Hopkins University. Participants were sorted into an experimental group, which received feedback for an extended period in the middle of the study, and a control group, which received no feedback until the very end. We hypothesized that members of the experimental group would perform better than the control group during the time that haptic feedback was administered, and that the performance improvement would remain after the feedback was removed. The results did not support the hypotheses; members of the control and experimental groups showed nearly identical learning curves.**

I. INTRODUCTION

Please note that I, Zachary Patterson, am submitting this paper as a draft to fulfill the NSF requirement for the Summer REU program. Because we are attempting to publish this work, this paper is the first draft. Also, we are still performing data analysis, so results are preliminary and subject to change.

Robotic Minimally Invasive Surgery (RMIS) is becoming more common in typical surgical procedures [1], [2], [3]. RMIS platforms such as the Intuitive Surgical da Vinci provide capability well beyond that of traditional laparoscopic surgeons and procedures thanks to better range of motion and dexterity. However, no commercially available, FDA approved RMIS platform provides haptic (tactile) cues to the user, presenting a major drawback. In traditional laparoscopic surgery, and most other forms of traditional surgery, surgeons use their sense of touch to interpret what they feel while manipulating patient tissue. The surgeon knows, intuitively from years of practice, how to react to these touch stimuli (e.g. the surgeon applies too much or too little force and reacts accordingly). Taking away the ability of the surgeon to feel within the patient, therefore, limits the procedures that can be performed with the da Vinci and causes a difficult learning curve.

The majority of research in this topic naturally focuses on producing a haptic feedback system for clinical use [4], [5], [6], [7]. However, there will most likely be a significant wait before any of these technologies are clinically or commercially viable. This leaves us with the important question of what to do in the interim to improve clinical practice of robotic surgery. The research described in this paper is a push to improve robotic surgery training by incorporating haptic force feedback. Previous work on the topic has shown that haptic feedback can help reduce user grip force [7], [8] and speed up learning for laparoscopic surgery [9]. In order to go further, to prove that haptic feedback improves training for surgical robots, we must show that haptic feedback not only improves performance, but also that the performance improvement remains after haptic feedback is taken away. By learning to associate force feedback with their vision in training, it seems likely that surgeons learning with haptic feedback would have an advantage and learn faster.

Therefore, in this paper we will test two hypotheses: *providing users with tactile feedback of forces produced during a training task will help them produce less force* and *the performance improvement will remain after haptic feedback is taken away*. We will use an inanimate training task, which are shown to correlate highly with *in vivo* training [10]. A previous paper by some of the same authors of this paper tested these hypotheses and found them to hold [11]. Much of the methods and apparatus are similar between that study and the study documented in this paper, with key differences in experimental design.

II. METHODS

Much of the setup and methods for this experiment were based on a study conducted by Brown et al [11]. Therefore, many aspects of this study are similar, if not identical.

A. Experimental Setup

For the experimental apparatus, we used an Intuitive Surgical da Vinci S and a tactile force-feedback system. The feedback system produces a squeezing sensation on the user's wrist. The amount that the system squeezes is proportional to the amount of force on the task board by the da Vinci tools. The system's components are as follows:

[1]Z. Patterson is with the Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, PA 15213 zachpatterson@pitt.edu

[2]B. Wolfinger is with the Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218 bwolfin1@jhu.edu

[3]K. J. Kuchenbecker is with the Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, PA 19104 and with the Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, Stuttgart, Germany kjk@is.mpg.de

[4]J. D. Brown is with the Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218 jdelainebrown@jhu.edu

a three axis force sensor on the task board, two wrist squeezing tactile actuators, signal conditioning circuitry, and a computer. Much of this system was originally designed for rating surgeon skill [12] and was later modified for this project.

The task board is a laser cut acrylic assembly that fits within the da Vinci skills dome. An ATI Mini40 SI 40-2 three-axis force sensor is mounted on the main platform of said assembly. The platform is held in place by magnets and dowel pins. The force sensor feeds into the signal conditioning circuitry through 6 differential analog inputs. The data is then processed in a series of filters, buffers, and analog to digital converters. This data retrieval and conditioning unit is controlled by a Teensy 3.1 microcontroller. The microcontroller then sends each sample to the computer through a USB serial interface. The computer then saves the data, scales it, and outputs the scaled signal to a wrist squeezing device described below.

The wrist-squeezing tactile actuator used was originally developed by Stanley and Kuchenbecker [13]. They found it to be highly intuitive for conveying the magnitude of stimuli [14]. They each consist of a Futaba s3114 servo motor mounted on a 3D printed frame. The servo motors have a maximum output torque of 0.17 Ns. The devices are secured to the wrist using a velcro strap that is looped through the 3D printed frame, around the wrist, and attached to a fixture on the motor. The servo is allowed to rotate through 50° from it's resting angle (Θ_{min}). As the angle of rotation of the servo increases, it tightens the strap, producing a squeezing stimulus on the user's wrist. The following equation demonstrates the angle of the servos:

$$\Theta_{\text{cmd}} = \begin{cases} \Theta_{\text{min}} & \text{if } F_{\text{mag}} < F_{\text{thresh}} \\ \Theta_{\text{min}} + \gamma F_{\text{mag}} \left(\Theta_{\text{max}} - \Theta_{\text{min}} \right) & \text{otherwise} \end{cases} \tag{1}$$

where Θ_{cmd} is the angle commanded to the tactile actuator's servo, $\Theta_{\text{min}} = 100^\circ$ is the servo's minimum allowed angle, $\Theta_{\text{max}} = 150^\circ$ is the servo's maximum allowed angle, F_{mag} is the time-varying magnitude of the three-axis force vector measured by the force sensor, F_{thresh} is a threshold for detecting non-zero force magnitude, and γ is a gain that adjusts the sensitivity of the feedback. $F_{\text{thresh}} = 0.1\,\text{N}$ and $\gamma = 0.714\,\text{N}^{-1}$ for this experiment.

A key addition to this experimental apparatus was the inclusion of "handedness" for the feedback; in other words, we wanted the subject to receive feedback to the hand that was actually causing the force. In the prior study, tactile feedback was output to both wrists regardless of which da Vinci tool was actually producing the force. In this study, handedness was included. In order to do this, conductive da Vinci tools (Maryland Forceps) were used. They were connected to a voltage dividing circuit board using the electrocautery connection on the tools. Additionally, the task board was grounded. Thus, when one of the tools was in contact with the task board, the voltage of that output pin greatly decreases and we can provide feedback to only the corresponding hand. The circuit was connected to the signal conditioning circuitry with a single-ended analog input for each hand.

The motor is position controlled by the PhidgetAdvancedServo 8-motor (1061_0) servo driver, which is in turn controlled by a python script running on the computer. The python script also records the data. The script was running at 50 Hz while subjects received haptic feedback and was running at 16.7 Hz when they did not. The subjects controlled the recording apparatus with a foot pedal.

B. *Experimental Protocol*

$N = 27$ subjects participated in this study (16 male, 11 female, mean age 24±10 years) from the general population of Johns Hopkins University. All procedures were approved by the Institutional Review Board under protocol # . Subjects were compensated with a $20 gift card.

The experiment followed a quasi-experimental design; all subjects completed the task identically, without haptic feedback, through the first part of the study. They were then sorted into groups based on their performance, which we measured using the mean force magnitude exerted by each participant, in order to minimize the difference in baseline skill between the groups. The groups are the Control Group (Group C), which received no haptic feedback until the eighth and final trial, and the Experimental Group (Group E), which received haptic feedback during trials three, four, and five.

After arriving at the testing site and sitting in front of the da Vinci console, subjects were asked to give informed consent. Then they were asked to complete a demographic questionnaire. The da Vinci system operation was explained. The participant was then given up to five minutes to complete a practice task to ensure familiarity with tool manipulation, clutching, and camera movement. After the practice task, the wrist-squeezing actuators were placed on the subjects and the haptic feedback was demonstrated. Subjects were shown how to use the recording pedal to properly start and stop recording. Finally, the task was explained using the task board as a visual. The selected task for the experiment is called the ring rollercoaster, chosen from the Intuitive Surgical da Vinci Skills Drill Practicum.

a) Ring Rollercoaster Task: The ring rollercoaster task is shown in Fig. 1. Participants move the ring from the left side of the metal track to the right side. No instructions were given about if, when, or how often to hand the ring from hand to hand. Instead of the standard rings, conductive O-Rings from Marco Rubber were used to allow conductive contact for the handedness functionality of the system (model number: S1104-010). Subjects were told to minimize the amount of force on the track (by trying not to touch it) while also completing the task as quickly as possible. They were also told to keep the ring towards the center of the camera view and to refrain from dropping the ring while transferring it between hands. Subjects completed eight ring rollercoaster trials. These trials are grouped into phases as shown in Fig. 2. After the first phase, subjects were sorted into groups. The

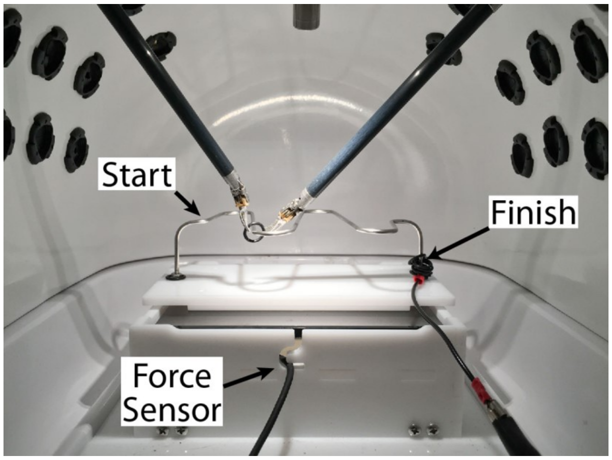

Fig. 1. Ring roller coaster task mounted on top of the task platform with embedded three-axis force sensor. One ring was placed at the start location for each trial. The ground circuit wire connects the task to the electrical circuit that detects left and right tool contact.

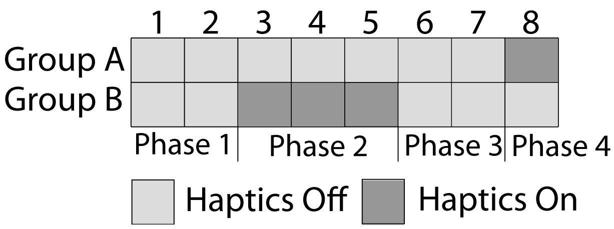

Fig. 2. Experimental protocol schematic showing when groups receive haptics and designating groupings (phases) of the trials.

Experimental Group received feedback during the second phase and the Control Group received feedback during the eighth phase. The tactile actuators were worn throughout all of the trials and white noise was played to isolate any audio cues and reduce distraction. After each phase, subjects completed a questionnaire about their subjective experience controlling the da Vinci during the previous phase.

Participants wore the tactile actuators for all nine trials, regardless of whether the haptic feedback system was active. Before participants began each trial, the tools were reloaded to reset their configuration, and the camera was adjusted to give a global view of the task board and the tool tips. After each phase, participants completed a questionnaire that contained quantitative and qualitative questions about their experience performing the task in that particular phase.

C. Metrics and Data Analysis

The force integral is used as the primary metric for evaluation. It takes into account both the magnitude of forces and duration of force application, giving us a metric that presents an aspect of both force and time.

1) Post-Phase Questionnaires: The questionnaire given to participants after each phase consisted of six questions taken from the NASA-TLX survey (ref) and additional questions. The latter questions are meant to capture the more specific subjective experience of the participants for this particular study. Participants answered each question on a 0-100 scale and had an optional text field for explanation. They were asked to keep explanations brief in order to keep the time of the study within the specified ninety minute window.

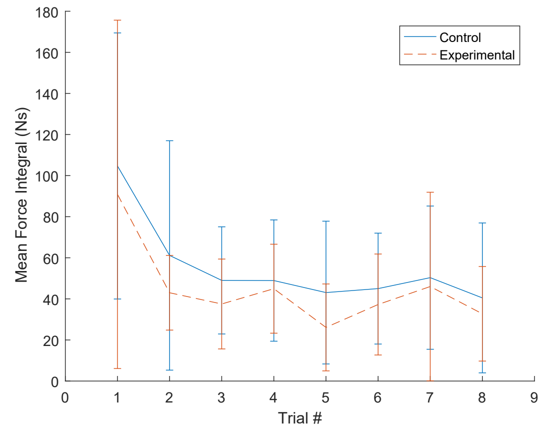

Fig. 3. Mean force integral for all participants in each trial. The dashed red lines indicate the experimental group, while the solid blue lines are for the control group. The error bars show the standard error of the mean.

2) Statistical Analysis: All statistical analyses were performed using R (v.3.3.2). A multilevel linear model was used to assess the effect of experiment phase and participant group on the two primary performance metrics. Within each model, participant was a random effect, experiment phase was a repeated-measure predictor, and group was a between-subject predictor. We determine significance using $\alpha = 0.05$

III. Results

The results of several subjects were discarded. One subject snapped several o-rings during the study. This subject did not follow directions to minimize force on the platform. Another subject was unable to complete the trial. A third and fourth subject were removed because of an error in the force feedback they received. This leaves the total participants to be analyzed at $N = 23$.

A. Force Integral

Fig. 3 shows the mean force integral for participants in each group for each trial of the experiment. Values for each trial were first averaged across participants before analysis. Clearly, one can see that there is no significant difference between the two groups ($\chi^2 = 1.03$, $p < 0.3095$). Contrasts were used to break down main effects of phase and group. The first contrast shows a significant difference in the force integral between phases one and two ($b = -36$, $t(63) = -3.30$, $p = 0.0016$, $r = .38$). The second contrast shows no significant difference between phases two and three ($b = 0.67$, $t(63) = 0.0616$, $p = 0.9511$, $r = .01$). The third contrast shows no significant difference between phases two and four ($b = -6.5$, $t(63) = -0.599$, $p = 0.5511$, $r = .08$). Three other contrasts were used to evaluate the main effect of group, but none of these were significant.

B. Survey

Survey questions were analyzed in a similar way to key metrics, looking at the main effects of phase and group. Few questions show notable and significant differences, and only those will be presented in this section.

IV. Discussion & Conclusions

The goal of this study was to show that a haptic force feedback system helps improve performance of novice users of the da Vinci surgical robot by reducing force exerted on the training platform and that this performance improvement remains after the haptic feedback is removed. Although a previous study's finding suggested this performance improvement, our study found that this finding may have been erroneous. In other words, haptic feedback did not cause the experimental group to perform significantly better than the control group (as measured by force exerted).

Although the hypotheses of this study were rejected, several topics of interest deserve discussion. First, why were the hypotheses not supported when identical hypotheses were supported in a similar previous study? Secondly, why didn't the haptic feedback improve performance? Literature discussed in the introduction suggests that it should. The discussion will deal with these questions and suggest future steps for this line of research.

To address the first question, we must compare the designs

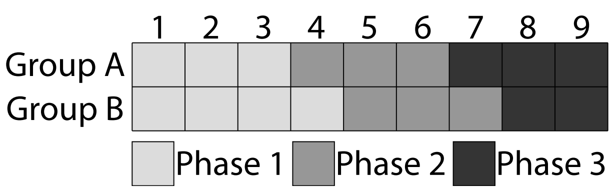

Fig. 4. Experimental protocol schematic for previous study showing when groups receive haptics and designating groupings (phases) of the trials.

of each of the studies. The previous study took place in 9 trials separated into three phases. In the second phase (a sequence of three trials), participants received haptic feedback similar to the feedback provided in the current study. In order to prove that the haptic feedback had an effect, participants were sorted into groups and phase two was introduced at different times between the two groups. See Fig. 4 for an illustration of this design. Staggering the feedback in this way was meant to show that improvement in phase two was a result of the feedback, and the results of the study suggested that this was the case. However, in the current study, a control group was compared to an experimental group. Both groups showed nearly identical trends in improvement. When we examine the results of the current study compared to the previous one, we see that the trends are similar. Fig. 5 shows the mean force integral from the previous study grouped by phase while Fig. 6 shows the same for the current study. The similarity of these trends suggests that a natural learning curve occurred in both studies and that the improvement seen in the previous study would have occurred whether or not haptics was

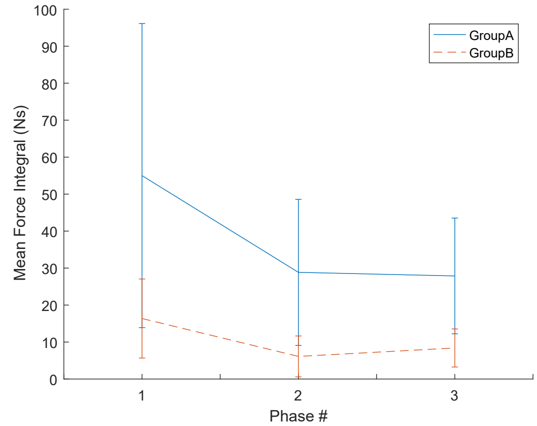

Fig. 5. Mean force integral for all participants in each experiment phase from the previous study. The dashed red lines indicate the experimental group, while the solid blue lines are for the control group. The error bars show the standard error of the mean.

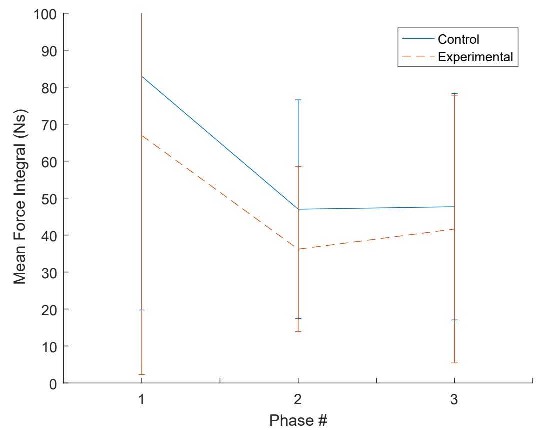

Fig. 6. Mean force integral for all participants in each experiment phase from the current study. The dashed red lines indicate the experimental group, while the solid blue lines are for the control group. The error bars show the standard error of the mean. Note that the fourth phase is truncated for better comparison to the previous study.

administered.

This still leaves our second question unanswered. Why didn't haptic feedback, at the very least, improve performance versus the control group? A number of potential explanations exist, many of them related. The first is that the difficulty of the ring rollercoaster task was not appropriate for haptic feedback. In particular, it is possible that the ring rollercoaster task was not difficult enough to perform. A user can clearly see whether or not the ring is contacting the track, so diligent users are able to perform the task well and improve quickly with only their vision. However, the ring rollercoaster task is not meant to simulate a surgical task, only to build hand-eye coordination and familiarity with the da Vinci. Specifically, force is to be minimized and avoided for the ring rollercoaster. In a task more similar to a surgical

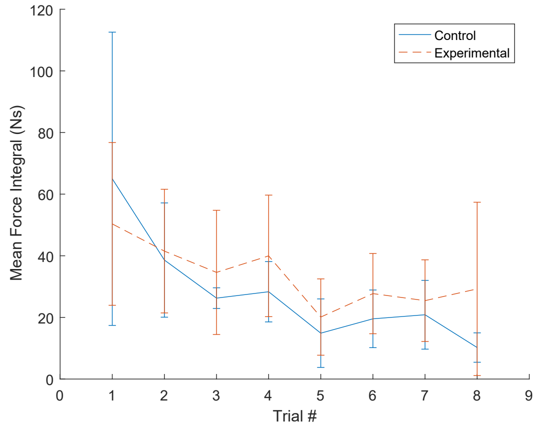

Fig. 7. Mean force integral for the ten "'Highest Skill"' participants in each experiment phase from the current study. The dashed red lines indicate the experimental group, while the solid blue lines are for the control group. The error bars show the standard error of the mean.

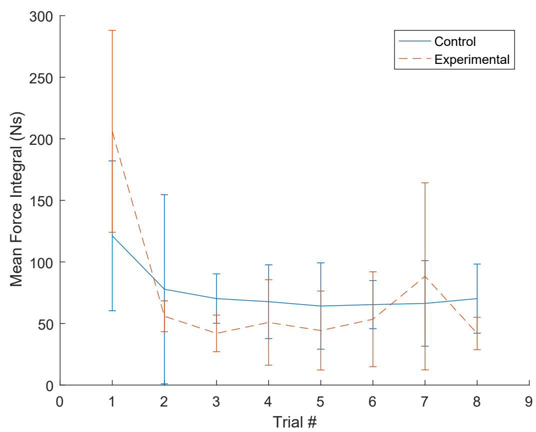

Fig. 8. Mean force integral for the ten "'Lowest Skill"' participants in each experiment phase from the current study. The dashed red lines indicate the experimental group, while the solid blue lines are for the control group. The error bars show the standard error of the mean.

task, where a certain amount of applied force is necessary for successful completion (e.g. suturing or dissection), haptic feedback might prove useful.

A related explanation is that certain users benefited from haptic feedback while others did not. Inherent user skill could play into this. If we take the ten most skilled and ten least skilled participants (as measured by force magnitude), it is clear that for the more skilled groups, haptic feedback did not improve performance whatsoever (see Fig. 7). In fact, some users of this set commented that the feedback was distracting. Conversely, when we examine the less skilled group, we see clear trends of improved performance for the experimental group (see Fig. 8). This sampling is not large enough to provide a concrete result, but it does point to a potential explanation: haptic feedback may be more useful to users who are less inclined to naturally learn to use the da Vinci.

REFERENCES

[1] K. K. Badani, S. Kaul, and M. Menon, "Evolution of robotic radical prostatectomy," *Cancer*, vol. 110, no. 9, pp. 1951–1958, 2007. [Online]. Available: http://doi.wiley.com/10.1002/cncr.23027

[2] A. L. Smith, K. M. Schneider, and P. D. Berens, "Survey of obstetrics and gynecology residents' training and opinions on robotic surgery," *Journal of Robotic Surgery*, vol. 4, no. 1, pp. 23–27, 2010. [Online]. Available: http://link.springer.com/10.1007/s11701-010-0176-0

[3] S. Maeso, M. Reza, J. A. Mayol, J. A. Blasco, M. Guerra, E. Andradas, and M. N. Plana, "Efficacy of the da Vinci surgical system in abdominal surgery compared with that of laparoscopy: a systematic review and meta-analysis." *Annals of Surgery*, vol. 252, no. 2, pp. 254–262, 2010.

[4] K. Bark, W. McMahan, A. Remington, J. Gewirtz, A. Wedmid, D. I. Lee, and K. J. Kuchenbecker, "In vivo validation of a system for haptic feedback of tool vibrations in robotic surgery." *Surgical Endoscopy and Other Interventional Techniques*, vol. 27, no. 2, pp. 656–64, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22806517

[5] A. M. Okamura, "Haptic Feedback in Robot-Assisted Minimally Invasive Surgery," *Current opinion in urology*, vol. 19, no. 1, pp. 102–107, 2009.

[6] L. Meli, C. Pacchierotti, and D. Prattichizzo, "Sensory subtraction in robot-assisted surgery: fingertip skin deformation feedback to ensure safety and improve transparency in bimanual haptic interaction." *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1318–27, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24658255

[7] C.-h. King, M. O. Culjat, M. L. Franco, J. W. Bisley, G. P. Carman, E. P. Dutson, and W. S. Grundfest, "A Multielement Tactile Feedback System for Robot-Assisted Minimally Invasive Surgery," *IEEE Transactions on Haptics*, vol. 2, no. 1, pp. 52–56, 2009. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/TOH.2008.19

[8] C.-h. King, M. O. Culjat, M. L. Franco, C. E. Lewis, E. P. Dutson, W. S. Grundfest, and J. W. Bisley, "Tactile Feedback Induces Reduced Grasping Force in Robot-Assisted Surgery," *IEEE Transactions on Haptics*, vol. 2, no. 2, pp. 103–110, 2009.

[9] M. Zhou, S. Tse, A. Derevianko, D. B. Jones, S. D. Schwaitzberg, and C. G. L. Cao, "Effect of haptic feedback in laparoscopic surgery skill acquisition," *Surgical Endoscopy and Other Interventional Techniques*, vol. 26, no. 4, pp. 1128–1134, 2012.

[10] A. J. Hung, I. S. Jayaratna, K. Teruya, M. M. Desai, I. S. Gill, and A. C. Goh, "Comparative assessment of three standardized robotic surgery training methods," *BJU International*, vol. 112, no. 6, pp. 864–871, 2013.

[11] J. B. Brown, J. N. Fernandez, S. P. Cohen, and K. J. Kuchenbecker, "A Wrist-Squeezing Force-Feedback System for Robotic Surgery Training," in *IEEE World Haptics Conference*, 2017.

[12] J. Brown, C. O'Brien, S. Leung, K. Dumon, D. Lee, and K. Kuchenbecker, "Using Contact Forces and Robot Arm Accelerations to Automatically Rate Surgeon Skill at Peg Transfer," *IEEE Transactions on Biomedical Engineering*, vol. PP, no. 99, 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7765041/

[13] A. Stanley and K. Kuchenbecker, "Design of body-grounded tactile actuators for playback of human physical contact," in *Proc. IEEE World Haptics Conference*, no. June, 2011, pp. 563–568. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5945547&tag=1

[14] A. A. Stanley and K. J. Kuchenbecker, "Evaluation of tactile feedback methods for wrist rotation guidance," *IEEE Transactions on Haptics*, vol. 5, no. 3, pp. 240–251, 2012.

APPENDIX

A. Research Ethics

I did an in depth training course in research ethics. Also, I had to work with the Johns Hopkins Homewood IRB to perform a human subjects experiment. Because of that, I had extensive exposure to research ethics for human subjects and am now fairly familiar with the responsible conduct of human research.

B. Value of the Program

Knowledge and skills gained during this experience will be instrumental for my future career. I plan on pursuing a PhD in robotics, so the skills obtained conducting a research project at a high powered robotics institution will translate well to that next step. In particular, the internship improved my technical skills, especially in programming and mechatronics. It also helped with "‘soft skills"’. I feel more confident designing experiments, writing papers, designing posters, and working with human subjects. Perhaps most importantly, since I led my own project (without a PhD mentor), I feel ready to manage a graduate level project throughout my PhD. If my research is in a similar field of research in grad school (human robot interaction), I could see this experience being even more helpful, as I've become familiar with the field over the course of the summer. Finally, the connections I've made during this experience have already proved valuable and I anticipate they will continue to do so.

C. Overview of the Program

Highlights:

- Interesting and cutting edge projects
- World class professors
- Potential for publishing
- Autonomous work (depending on your mentor, of course)

Areas for Improvement:

- The tours were hit or miss and should either be tailored to high skill engineering students or dropped from the program
- Short timeframe minimizes the potential impact of the projects