THE DEVELOPMENT AND VALIDATION OF A DYNAMIC HAPTIC ROBOTIC TRAINER FOR CENTRAL VENOUS CATHETERIZATION

A Thesis in
Industrial Engineering
by
Mary Yovanoff
© 2017 Mary Yovanoff

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science

August 2017
The thesis of Mary Yovanoff was reviewed and approved* by the following:

Scarlett Miller
Associate Professor of Engineering Design Industrial Engineering
Thesis Advisor

Jason Moore
Associate Professor of Mechanical and Nuclear Engineering

Janis Terpenny
Peter and Angela Dal Pezzo Department Head and Professor
Head of the Department

*Signatures are on file in the Graduate School
ABSTRACT

High-tech simulators are gaining popularity in surgical training programs due to their potential for improving clinical outcomes. However, most simulators are static in nature and only represent a single anatomical patient configuration. This is problematic because it limits a resident’s opportunity to practice adapting their skills to challenging but common patient cases. Virtual Reality (VR) simulation may have an advantage over these static simulators because they can present variations in patient anatomies. However, few studies have explored the impact of VR simulators on Central Venous Catheterization (CVC) procedures. This is important because 39% of patients who receive CVC’s experience adverse effects\(^1\) and these complications can be attributed to variations in patient anatomy\(^2\).

The purpose of this thesis was to design and develop a user interface and personalized learning system for a Dynamic Haptic Robotic Trainer (DHRT) that can present variations in anatomy and provide consistent objective feedback on performance, validate the DHRT system as an effective method for training the surgical skills necessary to place an Internal Jugular (IJ) CVC, explore the relationship between, self-ratings of performance, and subjectively measured ratings of performance and objectively-measured ratings of performance, and explore the impacts of the Dynamic Haptic Robotic Training system on CVC self-efficacy and CVC procedural skills when compared to current manikin based simulators. This was accomplished through four empirical studies conducted with a total of 56 medical professionals including: 26 medical students, 26 first year surgical residents and 4 expert surgeons. Specifically, two studies were conducted as part of this thesis during the design and development of the system. Feedback provided to novices during surgical skills training sessions was analyzed and used for the development of a graphical user interface and learning feedback system for the DHRT system. These studies also empirically and methodically identified 4 objective metrics for evaluating surgical skill proficiency in placing an IJ CVC. Two empirical studies explored the effectiveness of the DHRT system and manikin trainer for improving skill gains and self-efficacy in CVC insertion procedures.

Overall, this thesis work identified a methodology for systematically collecting and analyzing verbal and written feedback provided to novices during surgical skills training sessions and provides an empirical basis for the development of a graphical user interface and learning feedback system for the DHRT System for ultrasound-guided internal jugular catheterization based on this feedback. This thesis provides an empirical basis for the validation of a virtual reality \textit{haptic} training system for surgical procedures when paired with observer feedback, and contributes to the knowledge base of skill evaluation for IJ CVC placement. Additionally, this provides empirical evidence that showed that the DHRT and Personalized Learning Interface System is an effective method for improving \textit{self-efficacy} on CVC procedures, provides evidence that suggests surgical residents may not have a strong ability to accurately self-assess their performance, and provides empirical evidence that showed that the DHRT and Learning Interface System was an effective method for improving \textit{objectively measured skills} used in CVC procedures by providing direct, specific feedback.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... vii

LIST OF TABLES ............................................................................................................... x

ACKNOWLEDGEMENTS ................................................................................................. xi

Chapter 1 Introduction and Related Work .................................................................... 1

1.1 Central Venous Catheterization Insertion Procedures and Complications .... 4
1.2 The Use of Simulators in CVC Skills Training ................................................. 6
1.3 Current State of CVC Proficiency Evaluations and Opportunities for Advancement .............................................................................................................. 13
1.4 Research Objectives and Significance .............................................................. 16
   Objective 1 .................................................................................................................. 17
   Objective 2 ................................................................................................................ 17
   Objective 3 ............................................................................................................... 18
   Objective 4 ............................................................................................................... 18
1.5 Document Outline .............................................................................................. 18

Chapter 2 Development of the Dynamic Haptic Robotic Trainer Personalized Learning System .............................................................................................................. 20

   2.1 Phase 1: Systematic Analysis of Oral and Written Feedback during IJ CVC Training .................................................................................................................. 21
   2.2 Phase 2: Learning Interface Design and Development ................................. 25
   2.3 Phase 3: Usability Testing of the DHRT Personalized Learning System .......... 33
   2.4 Discussion ......................................................................................................... 36

Chapter 3 Comparing the Effectiveness of the DHRT System to Existing Manikin-Based Approaches ......................................................................................... 39

   3.1 Methodology ..................................................................................................... 39
   3.1.1 Participants .................................................................................................. 40
   3.1.2 Procedure .................................................................................................... 41
   3.1.3 Survey Instruments ................................................................................... 43
   3.2 Data Analysis and Results ............................................................................. 44
   RQ1: How, if at all, do medical students’ skill performance and CVC insertion self-efficacy change after training? ................................................................. 44
RQ2: Does the method of training impact gains in medical student CVC self-efficacy or CVC insertion skills? ................................................................. 45

3.3 Discussion.................................................................................... 46

Chapter 4 Can Haptic Simulators Distinguish Expert Performance? A Case Study in Central Venous Catheterization in Medical Education ....................................... 48

4.1 Methodology ............................................................................. 49
  4.1.1 Participants ........................................................................ 50
  4.1.2 Procedure .......................................................................... 50
  4.1.3 Metrics ............................................................................... 51

4.2 Data Analysis and Results ............................................................ 53

RQ1. Does the experience level of the participant (expert or novice) affect individual performance on the path length of the needle tip, the standard deviation of the deviations from an ideal path, the proximity of the needle tip to the center of the vein, the velocity of the needle tip, or the time required to complete the procedure when using the DHRT system during CVC training and controlling for patient case (one of the 8 anatomical variations)? ........................................... 54

RQ2. Can expertise be distinguished during CVC placement using the DHRT system through motion analysis? ........................................... 55

RQ3. Do novices improve their CVC insertion performance through use of the DHRT system? .............................................................. 57

4.3 Discussion.................................................................................... 59

Chapter 5 An Exploration into the Impact of Simulator Fidelity and Feedback on Surgical Resident Subjective Performance in Central Venous Catheterization Training ......................................................................................... 61

5.1 Methodology ............................................................................. 63
  5.1.1 Participants ........................................................................ 65
  5.1.2 Procedures .......................................................................... 65

5.2 Data Analysis and Results ............................................................ 72

RQ1: Do surgical residents improve their CVC self-efficacy over the course of training and is this improvement dependent on the type of training provided? ......................................................................................... 73

RQ2: How do perceived ratings of performance change throughout training and is this dependent on the training method? ........................................... 76

RQ3: How do perceived ratings of performance on the robotic trainer relate to objectively calculated performance scores in the DHRT system? ......................................................................................... 82

RQ4: How do perceptions of the effectiveness of the training simulator differ between simulator fidelity? .................................................. 84

5.3 Discussion.................................................................................... 86
Chapter 6 Measuring Surgical Residency Skill Gains in Central Venous Catheterization Training: An Exploration in Simulator Fidelity

6.1 Methodology
6.1.1 Participants and Procedures
6.1.2 Metrics
6.2 Data Analysis and Results
RQ1: How do subjective ratings of performance change throughout training and is this dependent on the type of training received?
RQ2: How do resident CVC skills quantified through motion analysis change through CVC training and is this dependent on the training method employed?
RQ3: Is there a relationship between objective and subjective measures of performance?

6.3 Discussion

Chapter 7 Conclusions and Future Work

Limitations and Future Work

References

Appendix A Patient Profiles
Vessel Placement for Initial 8 Patient Profiles
Final 17 Patient Profiles

Appendix B Surveys and Checklists
Central Line Self Efficacy Survey
Training Evaluation Surveys
IJ CVC Skills Checklist
Post-Test Checklist

Appendix C Training Scripts
Manikin Training Script
Using the Robot Script
Robot Training Script

Appendix D Coding Handbook for Feedback on Central Line Placement

Appendix E Self-Evaluation Training Book
LIST OF FIGURES

Figure 1-1: The Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) manikin which can be used with an ultrasound. It also has re-sealable “skin” which allows for repeated needle insertions and includes a hand-pumped arterial pulse..........................................................2

Figure 1-2: The syringe has been placed inside the Internal Jugular vein of the Blue Phantom Manikin Trainer with negative pressure applied through needle aspiration. The blue liquid present inside of the syringe in this image represents the ‘flash’ of blood that occurs when inside a vessel and aspirating on the syringe in a clinical patient ..............................................................3

Figure 1-3: The Dynamic Haptic Robotic System: During a practice needle insertion, the ultrasound probe is placed on the scanning surface and an ultrasound image will appear on the Monitor. When the needle is inserted into the scanning surface, the haptic robot will generate forces to simulate the feel of inserting a needle into a person..........................................................9

Figure 1-4: Retractable needle for the syringe. This allows the needle to appear as though it is inserting all the way into the scanning surface. The haptic robot will generate forces so that it feels like the needle is inserting as well. This allows for various profiles to feel different as the needle ‘passes through’ different types and thicknesses of tissues.........................................................10

Figure 1-5: The end of the syringe is equipped with a spring-loaded plunger. This allows the user to apply constant pressure, as though pulling the plunger out, which simulates aspiration - retracting the plunger to draw air or fluid into the syringe.................................................................11

Figure 1-6: When the Ultrasound Probe is placed on the scanning surface, an ultrasound image will appear on the monitor in front the user. In this figure, the ultrasound images is overlaid in the upper right corner.................................12

Figure 2-1: Frequency and type of feedback provided to participants during the pre- and post- test and how each item is addressed in the DHRT personalized learning system..................................................................................25

Figure 2-2: Home Screen for the DHRT system. At the top of the screen is a graph indicating performance over time, a personalized tip, and a high scoreboard. At the bottom of the screen are navigation buttons and a set of generalized tips. .........................................................................................27
Figure 2-3: Patient Profile Page: This page is presented prior to the start of each trial. On the left side of the screen is information about the patient that could be relevant to the procedure. On the right side are reminders for the procedure. .................................................................................................................. 28

Figure 2-4: Variations in Force for Profile 7: The amount of force required to insert a needle into the tissue will generally increase as the needle goes deeper in the tissue due to the friction forces between the needle and the tissue. How much the force increases depends on the type of tissue. For example, in this image there is a change in tissue type around 10cm and again around 15cm. .................................................................................................................. 29

Figure 2-5: Development of vessel positions for Profile 7: The text shown in the top image was used to identify key characteristics of this patient and a general outline for the ultrasound image was sketched. In this profile, the patient is obese which results in deeper vessels. The sketches were translated into the virtual profiles as shown in the bottom image.................................................. 30

Figure 2-6: Virtual Ultrasound Image: The needle tip is visible in the right vessel. The blue bar across the top of the screen indicates that the needle tip is in the vein. .................................................................................................................. 31

Figure 2-7: Personalized Learning Screen for the Dynamic Haptic Robotic Trainer. Each box provides feedback on (1) the overall performance, (2) the difficulty of the patient case, (3) the number of insertions and if the needle passed through the back wall of the vein, (4) the angle of insertion, (5) the distance to the center of the vein, and (6) the frequency of aspiration.................................. 33

Figure 2-8: Median responses on feedback survey items ............................................. 35

Figure 3-1: Needle paths of both an expert (blue) and a novice (red). The surface of the skin is a flat plane at Y = 0 (not pictured here). The blue tube represents the location of the vein and the red tube represents the location of the artery. The expert inserted slowly and then made micro-adjustments at the end of the insertion to precisely place the needle. The novice inserted the needle and missed the vein because they were lateral (too far in the positive x-direction), retracted the needle all the way out, and then inserted the needle again........................................................................................................................................ 52

Figure 3-2: (A) The triangle derived from needle path, insertion angle $\theta$ and vessel depth $H$. (B) Two needle paths with similar standard deviations but dissimilar standard deviation of the deviations. ................................................................. 53

Figure 3-3: Change in path length over time. Experts (top) showed no significant decrease or increase while novices (bottom) had a significant linear model for decrease in path length over time................................................................. 58
Figure 5-1: Practice Insertion Performance Evaluation Form included in Self-Evaluation Training Books ................................................................. 68

Figure 5-2: Ending a trial by placing the ultrasound probe in the red box. ............... 69

Figure 5-3: The Performance-rating Screen for the DHRT was presented to participants immediately after completing a trial but prior to receiving any feedback on their performance. .............................................................. 70

Figure 5-4: The Self-Evaluation Training Books were filled out by participants throughout the study. For each needle insertion, participants recorded the date, a self-rating of performance, the number of attempts required to access the vessel, any mistakes made, and if feedback was provided. ......................... 72

Figure 6-1: Attachment of the motion tracker (indicated by the red circle) to the syringe for the pre- and post- tests which was attached to the Ascension 3D Guidance trakSTAR electromagnetic position tracking system .................... 94

Figure 6-2: Plot used to identify the point of needle insertion. The Pythagorean Distance is shown in Blue and the Angle of the Needle tip is shown in Orange. In this case, data was cut at Time = 9.708 sec ........................................ 96

Figure 6-3: Participant performing a pre-test on the manikin and inserting the needle towards the head, which is the incorrect direction. .......................... 99
LIST OF TABLES

Table 3-1: Mean and Standard Deviation of Computed Metrics ........................................ 55
Table 3-2: Values of Binomial Logistic Regression Model .................................................. 57
Table 5-1: Self-Efficacy Responses ...................................................................................... 74
Table 5-2: Values of General Estimating Equation: Pre- and Post-SEL Survey Responses ................................................................................................................................. 75
Table 5-3: Values of General Estimating Equation: Pre- and Post-SEL Survey Responses ................................................................................................................................. 77
Table 5-4: Transcribed Errors from the Self-Evaluation Training Book by Participants AMON08 (Manikin) and CYTY07 (Robotic) and the Code of the Item From the Content Analysis ......................................................................................... 81
Table 5-5: Number of Errors Recorded by Each Training Group in the Self-Evaluation Training Book and the results of the Chi-Squared test. ................................. 82
Table 5-6: Results of the Mann-Whitney U-test which failed to indicate any significant differences in the training effectiveness between the Manikin and Robotic training conditions .............................................................................................................. 86
Table 6-1: Pre-Test Results for Participant IELE05 ............................................................... 100
Table 6-2: Results from One-Way Mixed ANOVA .................................................................. 102
Table 6-3: Values of Interaction Effect from 2-way Mixed ANOVA ..................................... 103
Table 6-4: Values of Main Effect for Test (Pre and Post) from 2-way Mixed ANOVA ................................................................................................................................. 103
Table 6-5: Values of Main Effect for Training Group from 2-way Mixed ANOVA ............. 104
Table 6-6: Example Scoring for A Pre- And Post- Test Of One Participant From Each Training Group ......................................................................................................................... 105
ACKNOWLEDGEMENTS

The research reported in this thesis was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL127316. It is important to note that the content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

I would like to thank my advisor, Dr. Scarlett Miller, for providing me with an opportunity to participate in such interesting and important research. She provided me with many occasions to grow both professionally and personally. I would like to thank David Pepley for his hard-work and dedication to this project as well as his patience and encouragement as a fellow grad-student. I would like to acknowledge David and Dr. Jason Moore, for their contributions to the physical development of the DHRT, as well as Dr. David Han, Katelin Mirkin, and Cassie Sonntag for their immense contributions to the medical knowledge required for this project, and for recruiting participants, and organizing space and equipment for all of the studies conducted for this thesis.

I would like to thank my fellow Britelab-ers: Elizabeth Starkey, Xuan Zheng, and Hong-En Chen for their regular encouragement and support. I also would like to acknowledge Beytullah Aydogan, Zachary Beard, and Kannan Kumar for their help in designing the user interface, my undergraduate assistants Andrew Strohmetz and Rucha Bhide for their valiant efforts to create a manikin from scratch, and Rohan Prabh for his help on the project. I would like to thank CAE Healthcare for loaning the Blue-Phantom trainer for use throughout my thesis research.

I am deeply and eternally grateful for my “somewheres”, who have loved, supported, and encouraged me through all of the ups and downs over the last two years. I truly do not know how I would have survived without all of you. Thanks for feeding me, zipping my coat when I broke my wrist, loaning me cats to snuggle with, playing board games with me, and being my family.
Chapter 1

Introduction and Related Work

More than 5 million Central Venous Catheters (CVC’s) are administered in the US each year\(^3\) in order to provide doctors with direct access to the heart for delivery of caustic or critical medications\(^4\). However, up to 39% of patients who receive CVC’s experience adverse effects\(^1\) and there is a high rate of morbidity in hospitalized patients\(^5,6\). Importantly, accidental arterial penetration, which occurs in up to 5% of adult patients and 26% of pediatric patients, is attributed to false identification of the target, spatial closeness of the artery and vein, or overshooting the final position\(^7\). In other words, variations in patient anatomy (e.g. skin thickness, amount of adipose tissue, vessel size and depth) make it significantly more difficult to place CVCs\(^8\) and thus contribute to the risk of mechanical complications during CVC placement\(^2\).

In order to reduce the surgical complications associated with CVC placement, surgical residency education has transformed from a “see one, do one, teach one” philosophy\(^9\) to a “see one, simulate many, do one competently, and teach everyone” model\(^10\) through the wide-spread integration of patient simulators\(^11\). These simulators have advantages such as increased standards for operating room efficiency\(^12\), decreased length of surgical clerkship\(^13\), improved ethics for practicing on real patients\(^14\), and an ability to provide a low-stress, no-risk method for training\(^15,16\). Simulators used in CVC training range from low fidelity homemade models\(^17,18\) to more “realistic” manikins that feature an arterial pulse (controlled through a hand-pump) and self-sealing veins (see Figure 1-1). While these more “realistic” simulators have “skin” that allow for the insertion of needles and allow for multiple practice trials without consequence\(^15\), they are static in nature and only represent a single anatomical configuration. Therefore, while residents are able to ‘simulate many’, they are simulating the exact same patient scenario. Because of this, it is not surprising that researchers have reported that while the most realistic CVC model can improve hand-eye coordination\(^19\) and resident comfort with CVC procedures\(^20\), they only temporarily improve procedural skills\(^21\). In addition, the benefit of this skill acquisition has been shown to decline over time\(^21\). This has led to large performance gaps between novice and expert surgeons\(^22\) and between completion of simulator training programs and surgical performance during the first several surgeries\(^23\).
These skill gaps may be due, in part, to the lack of variation of patient anatomy in these training systems. For example, research has indicated that including variation in scenarios that increase in difficulty when training motor skills increases skill transferability, skill retention and improves self-efficacy. However, no simulator to date has explored the implementation of these types of patient scenarios in CVC simulators.

In addition to a lack of variations in the type of patient scenarios provided by existing simulators, current simulators are also hindered by the type and timing of feedback provided during training. For example, the more realistic manikins currently used in CVC training only provide feedback through “blood flash” in the syringe if the resident has successfully punctured a vessel: blue for vein (target) and red for artery (see Figure 1-2). In order to receive any additional feedback on their performance, trainees must have a subjective observer present who can provide oral feedback, or feedback...
through a CVC checklist. This is problematic because these simulators not only rely on having an independent observer present during each training session, but the type and quantity of the feedback provided is not standardized during these sessions which can lead to large variations in the quality of feedback provided.

Figure 1-2: The syringe has been placed inside the Internal Jugular vein of the Blue Phantom Manikin Trainer with negative pressure applied through needle aspiration. The blue liquid present inside of the syringe in this image represents the ‘flash’ of blood that occurs when inside a vessel and aspirating on the syringe in a clinical patient.

In an effort to find more objective and immediate measures of performance for surgical training, other areas of surgery have begun to incorporate virtual reality (VR) simulators into training and have found them to be as effective and often more effective than their counterparts. This may be, in part, due to their ability to present realistic scenarios. In addition, optimally designed simulators can enhance learning by presenting increasingly difficult scenarios and providing personalized, objective feedback for each scenario. Importantly, preliminary research indicates that VR is effective for training ultrasound guided needle insertion procedures, but no research has been conducted to determine the effects of VR training on CVC insertion procedures.

In light of this prior work, The Dynamic Haptic Robotic Trainer (DHRT) VR system was developed to provide surgical residents with exposure to the effects of
variations in patient anatomy on CVC insertion using both visual and tactile feedback (see Figure 1-1). This Thesis discusses the development of the DHRT personalized learning interface and validation of the system for teaching CVC skills through experimental studies with medical students, surgical residents and expert surgeons. The remainder of this section highlights related research used to setup the theoretical foundation for this Thesis work including the steps of Central Venous Catheterization, state-of-the-art CVC skills training, the use of VR in medical education and the introduction of the Dynamic Haptic Robotic Training System. The chapter concludes with a statement of the research objectives and the outline of this thesis document.

1.1 Central Venous Catheterization Insertion Procedures and Complications

Central Venous Catheterization is a procedure conducted to provide medical personnel with direct access to a vein for the delivery of caustic or critical medications or when repeated access is needed. During CVC placement, an 18-gauge needle is inserted into a large caliber vessel, typically the internal jugular (IJ) vein or subclavian (SC) vein. Femoral catheters may be placed for temporary access during vascular or cardiac procedures but are removed after surgery. They are not typically used for long-term access due to a high risk of infection. The choice between IJ and SC lines is typically based on preference of the physician unless underlying anatomical structures indicate that one or the other is not suitable. For example, an IJ line may be more suitable when a patient’s anatomy is high risk for thrombosis (clotting of the vein) or pneumothorax (collapsed lung) because SC lines have higher complication rates for thrombosis pneumothorax and thrombosis. However, and SC line may be placed if the risk of arterial puncture could be life threatening for a patient, because IJ lines have higher rates of arterial puncture. SC lines may be recommended for long-term access due to lower infection rates although some literature suggests that there is not a difference in infection rates. Common complications for both IJ and SC lines include air embolisms (when and air bubble enters a vein), pneumothorax (collapsed lung), hemothorax (blood in the space between the chest wall and the lung), and arterial puncture which may be cause by multiple insertion attempts or inexperience. Literature is inconclusive on which is best and indicates that either IJ or SC may be appropriate, depending on the anatomy of the patient. However, research has shown
that internal jugular (IJ) catheters have the highest overall complication rates\textsuperscript{32} and, thus, is the focus of this thesis.

For IJ CVC placement, the use of ultrasound guidance is becoming standard practice because it can significantly reduce complication rates\textsuperscript{33}. During this procedure, surgeons use their non-dominant hand to move an ultrasound probe over the anatomical area and view the underlying structures (veins, arteries, etc.) on a monitor. Once a good view of the vessels is obtained on the ultrasound image, the needle is inserted at the apex of the triangle formed by the sternocleidomastoid muscle and the clavicle at a 30 – 45 degree angle while aspirating until flash is obtained, see Figure 1-3. Once the needle is in the center of the vein a guide wire is inserted followed by the placement of a hollow tube with control ports (catheter)\textsuperscript{41} with the catheter tip sitting close to the heart\textsuperscript{42}.

![Figure 1-3](image.png)

Figure 1-3: The triangle formed by the sternocleidomastoid muscle and the clavicle highlighted in blue and the Vein and the artery can be seen just underneath the muscles. Image modified from\textsuperscript{43}.

Problematically, more than 5 million CVCs are administered in the US each year\textsuperscript{3} and up to 39\% of patients who receive CVCs experience adverse effects\textsuperscript{1}. Importantly, a closed claims analysis found that a higher percentage of CVC complications were due to vascular access (the process described above) than maintenance issues and that nearly half of the reported complications were preventable\textsuperscript{44}.
Due to the close proximity of the arterial jugular vein and the carotid artery, arterial puncture was the most common mechanical (noninfectious) complication in IJ CVC placements prior to the use of ultrasound guidance for placement\textsuperscript{37,45}. While the overall complication rate has been reduced through the use of ultrasound guidance\textsuperscript{46}, arterial puncture is still the most frequent mechanical complication\textsuperscript{47}. Arterial puncture is problematic because it can cause a hematoma (excessive bleeding inside the body), but further complications can arise if medical personnel are unaware that they have punctured the artery, rather than the vein, and proceed to place the catheter\textsuperscript{48}. Simulators have been introduced into medical education in an effort to reduce these complications\textsuperscript{49} and improve patient outcomes during CVC procedures.

### 1.2 The Use of Simulators in CVC Skills Training

Surgical residents have traditionally been trained using the Halstedian apprenticeship model of “see one, do one, teach one”\textsuperscript{9}. However, over the last 20 years this method has largely been replaced by virtual patient simulators\textsuperscript{11} due to stresses on the apprenticeship model such as: increased standards for operating room efficiency\textsuperscript{12}, decreased length of surgical clerkship\textsuperscript{13}, and the ethics of practicing on real patients\textsuperscript{14}. These simulators have been praised for their ability to provide a low-stress, no-risk method for training\textsuperscript{15,16} and their potential to transform medical curriculum from a “see one, do one, teach one” model to a “see one, simulate many, do one competently, and teach everyone” model\textsuperscript{10}. Importantly, the use of medical training simulators has been shown to reduce complication rates\textsuperscript{50}.

The current state of the art in CVC training includes a “realistic” Simulab CentralLineMan manikin that features an arterial pulse (controlled through a hand-pump) and self-sealing veins (see Figure 1-1). While these “realistic” simulators have “skin” that allows needles to be inserted and allows for multiple practice trials without consequence, they are static in nature and only represent a single anatomical configuration. Therefore, while residents are able to ‘simulate many’, they are simulating the exact same patient scenario. While practice is important to initially learn a skill, repeatedly practicing the same skill in the same scenario may not increase skills past a basic proficiency level\textsuperscript{51}. This is because research has shown that the learning process is optimized when learners are given objectives that are scaled to their proficiency level\textsuperscript{30}. 

---

6
Importantly, research has shown that increasing the difficulty of a task throughout the learning process helps refine skills and attune the learner to more subtly difficult aspects of the motor skill. Thus, introducing variability into learning systems can increase skill retention when later tested on similar variations of the initial task and can result in increased task performance when compared to practicing the exact same scenario. This is important because research has suggested that the relationship between self-efficacy (confidence in one’s ability to perform a skill) and performance may decrease when applying the task to novel situations because individuals may not be able to accurately predict the skills necessary to complete the novel task. For example, research investigating self-efficacy in endoscopic skills found that participants who were part of a practice only intervention (compared to self-observation or expert observation) had a decreased ability to accurately judge their performance. Research has also suggested that for novice trainees there is no correlation between self-efficacy and performance on non-technical skills, such as situation awareness and decision-making. Additionally, when high performance is expected, skill performance may break down despite high self-efficacy; in other words, individuals may “choke under pressure.” When taken together, this research suggests using a static manikin that presents only single anatomical scenario, no matter how realistic, may not be optimal for training medical professionals in CVC placement. Thus, it is not surprising, that researchers have reported that while the most realistic CVC model can improve eye-hand coordination and resident comfort with CVC procedures, they only temporarily improve procedural skills and the benefit of this skill acquisition has been shown to decline over time.

Due to the deficiencies of current simulation trainers, researchers have begun to explore the use of Virtual Reality (VR) as a method for training surgical procedures because it provides a realistic environment to learn and practice surgical skills without compromising patient safety. While VR models have not yet penetrated the CVC curriculum there has been research that has shown promise for the use of VR systems in surgical education in areas such as cysto-urethorscopy, radiology, and laparoscopy. For example, a recent study that examined the transfer of skills from virtual reality training to clinical procedures showed that participants performed better than their counterparts on both the specifically trained skills, such as palpating the patient, as well as generic skills used in clinical situations. In addition, a study conducted on laparoscopic cholecystectomy training found that participants who were trained with the virtual reality simulator had superior technical performance post-intervention compared
to their counterparts who received standard training\textsuperscript{61}. Another study comparing VR training to standard training in laparoscopic surgery found VR to be more effective at reducing errors when transferring to real world applications\textsuperscript{60}. The use of VR for CVC training is supported by studies which found VR simulators provide sufficient realism for teaching ultrasound guided needle insertion procedures\textsuperscript{31}. Specifically, studies have suggested that use of augmented reality (AR) is feasible in CVC training\textsuperscript{62}.

In light of this prior work, the Dynamic Haptic Robotic Trainer (DHRT) VR system was developed to teach CVC training by presenting variations in patient anatomy using both visual and tactile feedback (see Figure 1-3). The use of this type of VR simulator in surgical training may have a major advantage over standard medical simulators because VR systems can provide real time tactile and visual feedback that correspond to variations in anatomy and therefore present multiple realistically difficult situations\textsuperscript{29}. 
The DHRT consists of a 3D Systems Geomagic Touch X (Rock Hill, SC), a virtual ultrasound system utilizing an Ascension 3D Guidance trakSTAR (Sheldburne, VT) electromagnetic position tracking system and a 3D printed ultrasound probe, and software visualizations and recordings developed in MatLab and Simulink. For details see Pepley et. al\textsuperscript{63}. The Geomagic X provides positional data to the simulation, as well as haptic feedback to the user based on needle insertion characterizations\textsuperscript{64}. Importantly, the needle tip on the syringe retracts (see Figure 1-4) so that only the first few millimeters enter the scanning surface, which ensures that all forces on the needle tip are created by the haptic robot. Additionally, the plunger on the syringe is connected to a spring so that continuous force can be applied to simulate aspiration (see Figure 1-5). When the

Figure 1-3: The Dynamic Haptic Robotic System: During a practice needle insertion, the ultrasound probe is placed on the scanning surface and an ultrasound image will appear on the Monitor. When the needle is inserted into the scanning surface, the haptic robot will generate forces to simulate the feel of inserting a needle into a person.
ultrasound probe is placed on the scanning surface an ultrasound image appears and responds as an actual ultrasound would (see Figure 1-6).

Figure 1-4: Retractable needle for the syringe. This allows the needle to appear as though it is inserting all the way into the scanning surface. The haptic robot will generate forces so that it feels like the needle is inserting as well. This allows for various profiles to feel different as the needle ‘passes through’ different types and thicknesses of tissues.
Figure 1-5: The end of the syringe is equipped with a spring-loaded plunger. This allows the user to apply constant pressure, as though pulling the plunger out, which simulates aspiration - retracting the plunger to draw air or fluid into the syringe.
The DHRT trainer has the potential to advance CVC medical training beyond the 100 year-old Halstedian apprenticeship model and beyond traditional static simulators due to its ability to provide real time feedback which can improve the efficiency of skill acquisition and its ability to provide a wide range of anatomical scenarios to train surgeons on complications associated with patient variability. However, this method of haptic training has not been validated for its ability to train surgical residents on CVC skills or identify if, or to what effect, it can help surgical residents reach expert performance. In addition, the initial development of the DHRT system did not have a personalized feedback system that denoted resident’s performance on the trainer reducing its utility for improving CVC skill gains.

Figure 1-6: When the Ultrasound Probe is placed on the scanning surface, an ultrasound image will appear on the monitor in front the user. In this figure, the ultrasound images is overlaid in the upper right corner.
1.3 Current State of CVC Proficiency Evaluations and Opportunities for Advancement

It is important to consider the evaluation system, in addition to the training system, used to determine if a surgical resident is ‘proficient’ at a surgical skill. Importantly, there is no single standard for evaluating skill competency in CVC placement. However, most studies evaluating CVC skills training make use of various types of checklists including procedure specific self-evaluations\textsuperscript{66}, and supervisor-evaluations\textsuperscript{27}. It has been suggested that the evaluation of competency of a trainee at a surgical procedure should include a combination of the trainee’s self-perception of ability, an evaluator’s perception of their ability, and an objective skills based evaluation of their ability\textsuperscript{67}. This section serves to highlight these evaluation methods, how they relate to expertise development, and future opportunities for VR-based simulator feedback.

One of the most widely-integrated methods for assessing CVC proficiency is through the use of binary ‘check-lists’ completed by an independent observer. Specifically, during these testing scenarios, the surgical resident thinks out loud while demonstrating the CVC insertion procedure on a manikin. A senior resident or attending observes their performance and ‘checks’ off each item on the binary checklist. These checklists are used to compare trainee skills to some baseline and evaluate if they are good enough to ‘pass’ and perform a the procedure in a clinical scenario\textsuperscript{68,69}. However, there is no ‘gold standard’ checklist used across all residency programs which has led to a wide variety of CVC skills check lists such as the Global Rating Scale (GRS) Assessment tool that uses a 9-question Likert scale assessment intended for use generally across multiple skills\textsuperscript{70}, a 29 item binary checklist published as a valid method for evaluating final CVC skills\textsuperscript{71}, and checklists that integrates both cognitive and procedural steps of the CVC procedure\textsuperscript{51,72}.

One of the biggest problems with these types of checklists is that they encourage “teaching to the test”\textsuperscript{73} rather than teaching a specific core set of skills. In addition, checklists can result in observers simply checking off boxes rather than providing helpful feedback\textsuperscript{74}. This is problematic because providing detailed and appropriate feedback is a critical part of the learning process and the opportunity for deliberate practice to incorporate that feedback is crucial to the development of expertise in surgery\textsuperscript{75}. In fact, research has shown that the timing and type of feedback (task specific, processes specific, or self-regulatory) provided to an individual can change what and how they learn\textsuperscript{26}. This
suggests the need to distinguish between evaluation tools and training tools\textsuperscript{76}. There has been some effort to develop tools more appropriate for training rather than just skill evaluation\textsuperscript{77}.

One way to combat the deficits of these existing testing methodologies is through the integration of VR systems in surgical education because these systems can provide objective, specific, personalized feedback for every practice attempt. If well-designed, VR systems can address all three necessary stages of surgical skill acquisition: cognitive (knowledge), associative (technical skill), and autonomous (adequate judgement) skills\textsuperscript{78}. Importantly, VR systems can be designed to present personalized feedback in both textual and graphical forms which may additionally improve information retention\textsuperscript{79}. Thus, VR systems that are thoughtfully designed and integrated into surgical training programs could lead to improved gains and transferability of skills\textsuperscript{80}. If designed to effectively and objectively capture skill performance, VR systems may be a feasible training system in surgical fields as well as a tool for skill evaluation\textsuperscript{81}. However, before an effective VR Feedback system can be designed, it is first important to consider what denotes appropriate feedback in these types of training environments.

Importantly, research has shown that surgical expertise is developed through deliberate practice and feedback; and the opportunity to incorporate the feedback is critical for developing expertise in surgery\textsuperscript{75}. Specifically, research has suggested that, when learning motor tasks, it may be more beneficial to give individuals feedback on their performance after completing successful trials instead of poor trials, give corrective feedback about small errors rather than large ones\textsuperscript{82-84}, and provide positive feedback over consistent negative feedback\textsuperscript{85}. In addition, this research has shown that learning is enhanced when the skills are presented as learnable rather than an inherently fixed skills\textsuperscript{86} and feedback is given such that the learner can immediately correct small mistakes so that confidence, or self-efficacy, is built gradually\textsuperscript{25}. Providing feedback after good or accurate performance increases both skill performance and self-efficacy separately\textsuperscript{82,87} and additional research shows that increased self-efficacy directly impacts skill performance\textsuperscript{54,88} and motor skill acquisition\textsuperscript{89}. When learning a relatively simple motor skill, learners may be cognitively lazy and complete several repetitions of the skill with only minimal cognitive effort which may inhibit memory consolidation and, therefore, skill retention\textsuperscript{90,91}. Memory consolidation can be enhanced by encouraging learners to consider future situations in which they would expect to use the skills they are learning\textsuperscript{92}. 
Additionally, providing learners with augmented feedback (feedback not normally present) can help engage cognition during the learning process.\(^93\)

The link between positive self-efficacy and increased task performance may be the result of greater attention allocation to the relevant task.\(^{94}\) This is supported by research in CVC training which has shown that, individuals who were asked to identify and discuss errors in performance while learning the procedure had better skill retention than their counterparts who simply learned the skills only.\(^{95}\) Research has shown that receiving feedback on cues that indicate performance, such as identifying errors, can help a learner be more accurate as assessing self-performance.\(^{96}\) This has large implication for feedback during surgical training because research has indicated that there is little or no relationship between self-assessment and objective measures of performance and, worse, individuals with lower objective scores of performance are worse at self-assessing their performance.\(^{97}\) For example, in a knee joint injection, confidence was inversely related to objective measures of performance but was directly related self-assessed performance. In other words, participants were over confident in their ability to perform the task and were also inaccurate in assessing their own performance on the skill.\(^{98}\) Some research has shown that, while the correlation between self-efficacy and actual performance of surgical skills is weak, it does improve with experience.\(^{99}\) These studies indicate the importance of the type and quality of the feedback given for helping novices accurately assess their own performance.

However, before an appropriate feedback system can be developed for CVC insertion procedures, it is first vital to determine the key metrics that distinguish expert performance. One such method for operationalizing expertise is through motion analysis of hand movements, which has been validated as a means of distinguishing experts from novices in surgical procedures\(^{100,101}\) with technology such as the Imperial College Surgical Assessment device (ICSAD)\(^ {99,102}\). Importantly, Kim et al.\(^{103}\) looked at motion analysis specifically in CVC placement and Varas et al.\(^{104}\) validated motion tracking as means of assessing movement in simulated CVC placement. However, it is still important to denote what aspects of motion analysis distinguishes this expertise. While not studied in CVC procedures, a study in Arthroscopy found that experts had a significantly shorter path length than novices\(^ {105}\). A study in laparoscopic surgery also found that experts have a shorter overall path length for the movement of their tools\(^ {106}\). Another study validated the use of electromagnetic motion tracking to distinguish between expert and novice performance by looking at path length, translational movement, and rotational movement.
of the needle hand as well as movement of the ultrasound hand finding that experts had significantly less hand movement\textsuperscript{107}. However, path length of the needle tip has not been studied for any procedure involving needle insertions. In addition to evaluating metrics for motion analysis, all of these studies noted that experts completed the task in a shorter amount of time than novices. Interestingly, the precise type of movement seems to depend on the task being performed. For example, in laparoscopy, where the task involves cutting a smooth line, experts have a smoother path than novices\textsuperscript{107}, but in endoscopic knot tying experts had a wider range of motion characterized by rapid controlled motions\textsuperscript{108} when compared to novices.

One method of evaluating smoothness of motion during surgery is jerk, which is the derivative of acceleration with respect to time\textsuperscript{109}. These metrics (time to complete, path length, velocity of movement, and smoothness of motion) can be captured using VR simulators\textsuperscript{110,111}. Importantly, the use of electromagnetic sensors to identify motion paths during CVC training show promise as a valid and objective assessment of skills\textsuperscript{112} and these metrics can be captured using VR simulators\textsuperscript{113}. However, the utility and appropriateness of these metrics for distinguishing expertise in CVC insertion has yet to be evaluated and therefore it is unclear if, or how, feedback on these skills should be integrated into new high-fidelity CVC simulators like the DHRT system.

\textbf{1.4 Research Objectives and Significance}

The literature presented in this chapter provides evidence that there is a large skill gap for novice surgeons when transferring skills from simulator training to clinical situations for Internal Jugular Central Venous Cather (IJ CVC) placement\textsuperscript{23}, and this can lead to severe complications sometimes resulting in the death of the patient\textsuperscript{5,6}. This skill gap may be due to a lack of exposure to anatomical variation during training\textsuperscript{24} and insufficient feedback on skill performance\textsuperscript{74}. Research has indicated that providing exposure to variations of a scenario during motor skills training can increase the effectiveness of skill transfer to novel situations\textsuperscript{24} but this concept has not been explored in the context of CVC placement. Additionally, research has indicated that Virtual Reality simulators can be effective methods for training and evaluating surgical skills\textsuperscript{31}, but there has been limited research on the effectiveness of a virtual reality trainer designed specifically for CVC placement.
In order to address these research gaps, the main purpose of the research presented in this thesis was to design a Dynamic Haptic Robotic Trainer (DHRT) and learning feedback system for training ultrasound-guided internal jugular catheterization, identify objective metrics for evaluating CVC skills competency, and to explore the effectiveness of the DHRT system for training surgical residents on CVC placement. Specifically, this thesis aims to address the following 3 research objectives:

**Objective 1**

Design and develop a user interface and personalized learning system for the Dynamic Haptic Robotic Trainer that can present variations in anatomy and provide consistent objective feedback on performance in order to understand how the DHRT training system can increase knowledge retention and increase effective skill transfer from practice on a simulator to performance on clinical patients. Research indicates that Virtual Reality (VR) simulators can effectively train surgical skills\(^{28}\), therefore the development of VR simulator specific to CVC placement is the first step towards understanding the impacts of using VR for CVC training.

**Objective 2**

Validate the DHRT system as an effective method for training the surgical skills necessary to perform IJ CVCs and utilize motion tracking to identify objective measures for evaluating procedural competency. While research has shown that motion tracking can be an effective method for evaluating skills competency in surgical procedures\(^{100,101}\), there is limited research in exactly what metrics exemplify expertise in CVC procedures. Therefore, research that identifies specific metrics for evaluating CVC skills can contribute to the effective validation of CVC training systems and evaluation of CVC skills competency.
Objective 3

Explore the relationship between, self-ratings of performance, and subjectively measured ratings of performance and objectively measured ratings of performance. Research has shown that there is a direct relationship between self-efficacy and skill performance \(^{54,88}\), but that there may not be a strong relationship between self-ratings of performance and actual performance \(^{98}\). However, training novices to identify and correctly interpret cues that indicate adequacy of performance through appropriate feedback can help improve this relationship \(^{96}\). Therefore, research that examines the relationship of self-ratings of performance, observer evaluated performance and objectively measured performance can help identify if residents are able to accurately assess their ability to perform a skill.

Objective 4

Explore the impacts of the Dynamic Haptic Robotic Training system on CVC self-efficacy and CVC procedural skills when compared to current manikin based simulators. The use of manikin simulators for CVC skills training can improve skills training \(^{19}\) and resident comfort with procedures \(^{20}\), but it remains how skills and self-efficacy are impacted by the use of the DHRT for CVC training. Therefore, research that provides empirical evidence on the impact of DHRT training on CVC skills and self-efficacy can help researchers identify how to best utilize Virtual Reality technology to train CVC skills and ultimately reduce complications during CVC placement.

1.5 Document Outline

In order to address the research objectives of this thesis, this document presents a total of 7 chapters detailing the development and validation of the DHRT system and the exploration of the impacts of DHRT training on self-efficacy and skill gains for central venous catheter (CVC) procedures. The first chapter of this thesis provided a detailed literature review and served to familiarize the reader with CVC procedures and the
associated complications as well as current training methods and support for improved training methods using virtual reality simulation and consistent objective feedback.

Chapter 2 describes the design and development of the DHRT personalized learning system and discusses the results and modifications made to the system after user-testing. Chapter 3 presents findings from a study that was conducted on 18 third-year medical students to determine if the DHRT was effectively training surgical skills when feedback was provided by an observer. Chapter 4 presents a study which compared expert performance to novice performance on the system and identified objective metrics for distinguishing between the two groups. Chapters 5 and 6 present the results of a study conducted with 26 first-year surgical residents at Hershey Medical Center. Half of the participants were trained using the DHRT system and received feedback through the personalized interface and half were trained using the manikin simulator and received feedback from an observer. Chapter 5 discusses the impacts of these training conditions on self-efficacy and self-rating of performance while Chapter 6 discusses the differences in both objectively and subjectively measured skill gains. Chapter 7 serves as a summary of the work conducted including the limitations and future direction of the research.
Chapter 2
Development of the Dynamic Haptic Robotic Trainer Personalized Learning System

Publication: Mary Yovanoff, David Pepley, Katelin Mirkin, Jason Moore, David Han and Scarlett Miller. Personalized Learning in Medical Education: Designing a User Interface for a Dynamic Haptic Robotic Trainer for Central Venous Catheterization. To Appear in the Human Factors and Ergonomics Society Annual Meeting, October 9–13, 2017, San Antonio, TX.

The DHRT system was developed with the goal of increasing CVC needle insertion performance, particularly as it pertains to variations in patient anatomy, and skill transfer from simulated training to clinical patients in the operating suite. The purpose of this chapter is to discuss the methodology and results of a user study that was conducted in three phases to: (1) systematically analyze the verbal and written feedback given to individuals being tested on IJ CVC procedures (2) incorporate the information gathered in Phase 1 into a personalized learning interface for the DHRT system and (3) evaluate the initial design of the learning interface through user testing and interviews. This is important because while current research has shown that high-fidelity simulators are more effective at reducing clinical errors compared to training without a simulator at all\textsuperscript{49}, research has also shown that there is a large increase in skill gains during the first several surgeries on clinical patients\textsuperscript{23}. This may be due, in part, to the quality and type of feedback received during the training process which typically consists of subjective verbal feedback and a checklist\textsuperscript{27}. Feedback is critical for learning but the delivery and quality of feedback can change its effectiveness\textsuperscript{26}. The criteria for evaluating skills and the means in which the information is displayed must be appropriately designed and implemented to ensure the feedback is appropriately understood and integrated\textsuperscript{105}. The remainder of this chapter describes the design, development and testing of the DHRT Graphical User Interface (GUI) personalized learning system.
2.1 Phase 1: Systematic Analysis of Oral and Written Feedback during IJ CVC Training

The purpose of Phase 1 was to systematically analyze the feedback given to individuals being tested on IJ CVC procedures. Specifically, this was done through an experimental study with 18 medical students (11 male, 7 female) between the ages of 23 and 35 with a mean age of 26. The details of the study follow.

2.1.1 Procedure

At the beginning of the study, the purpose and procedures were explained, any questions were answered, and informed consent was obtained for participation in the study and use of video and audio recording equipment. All participants consented to the use of video and audio recording and thus the entirety of the study for all participants was recorded.

Groups of three participants were given a demonstration of central-line placement from a second-year medical resident with expertise in the area using a Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) manikin (see Figure 1-1). During the training session, participants were shown how to use an ultrasound, how to identify and distinguish between the artery and vein, how to use anatomical landmarks as a guidance for line placement, how to insert a needle, how to identify needle location based off of ultrasound feedback and how to confirm needle placement using flash feedback. Importantly, the CVC training procedures employed in the current study are the same procedures used to train individuals in the surgical residency program at HMC and the second-year resident who performed the training in the study had conducted numerous CVC training sessions at HMC.

Once the training was complete, each of the participants individually entered into a pre-test where they were asked to insert a needle for central line placement into the same Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) used for the demonstration while verbalizing what they were doing according to a think-aloud procedure\textsuperscript{114}. During the pre-test, each participant was observed and evaluated by the same second year medical resident using a modified Internal Jugular Catheterization (IJ CVC) evaluation form. The IJ CVC evaluation form is a 10-item checklist focusing
exclusively on the needle insertion portion of the procedure and including items like “continuously aspirating the entire time”, “conducting the entire procedure without any mistakes”, and “selecting the appropriate site for venipuncture” (see Appendix B for full checklist). Participants were not provided any feedback throughout their pre-test, but once the pre-test was complete, they were informed if they had successfully placed the needle and what errors occurred during the procedure.

After receiving feedback, each of the participants individually participated in one of three training methods as follows:

*Manikin training*. Participants in this condition performed all of their training on a second Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660).

*Robotic training*. Participants in this condition performed all of their training on the DHRT.

*Mixed training*. Participants in this condition performed half of their training (4 patient profiles) on the Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) and half of their training on the DHRT.

For the training conditions, participants were provided with 1 of the 8 original patient profiles developed by the team. Once the participant read the profile they were asked to place a needle in the center of the vein while using the think-aloud procedure. Unlike the pre-test, a researcher gave verbal corrections as participants attempted the procedure. After the participant successfully entered the vein or accidentally hit the artery they were provided with verbal feedback on their performance (e.g. the final needle position, average insertion angle, number of attempts and insertion technique) and were able to ask any questions they had. Once complete, a new patient profile was provided and this procedure continued until all 8 profiles were completed. This training took approximately 20 minutes.

Once participants had completed all 8 practice trials, they individually completed a post-test assessment where they were asked to perform the CVC insertion on a Kyoto CVC Insertion Simulator II (Model #M93UB). This manikin was selected for the post-test because it differed in anatomical structure (e.g. depth and size of vessels, skin thickness, and vessel wall thickness) from both the pre-test and training manikin. The
same procedures from the pre-test were used for the post-test and the same medical resident conducted this examination. After the test was complete, the participant was given extensive feedback on their performance and given the opportunity to see a demonstration of the rest of the central line placement procedure (e.g. insertion of the guide wire and catheter placement).

2.1.2 Data Analysis

In order to identify the feedback being provided during CVC training, a qualitative analysis was conducted on the feedback provided during the pre- and post-tests. Specifically, the video recordings of the pre- and post-test were transcribed and analyzed by two independent raters using combined principles of inductive and deductive content analysis\textsuperscript{115} in NVivo v11.4.0. Because a pre-existing coding scheme did not exist, the Internal Jugular Catheterization (IJ CVC) evaluation form (See Appendix B), used to evaluate resident performance on the pre- and post-tests, was used as a starting point to develop a coding scheme to analyze the verbal feedback provided to the participants after testing.

During transcript coding, each item on the checklist was made into a node in NVivo. Next, two researchers verbally discussed each item on the checklist until both felt satisfied that they had a mutual understanding of each item. For example, “inserting needle at a 30-45 degree angle” was coded any time the specific angle of the needle was mentioned or any time “shallow” or “steep” was used to describe the needle. Each rater coded 3-5 different transcriptions and were encouraged to add any nodes they felt were frequently mentioned. An example is “hand position on the syringe”. Together the raters then reevaluated the existing nodes to form a consensus for which nodes were a good representation of the feedback being provided. A coding handbook was developed which can be found in Appendix D.

After thoroughly discussing the handbook, the raters both individually coded all eighteen transcriptions (one transcription contained both the pre- and post-test for a participant). An example portion of a transcription and accompanying code is shown below: the first highlighted portion was coded as “aspirating the entire time” while the second highlighted portion was coded as “identifying the vessels on the ultrasound”.
Participant: “Do I need to inject this?”
Evaluator: “No you’re fine. You can just leave that right there. That’s fine. Um, Excellent job. I’ll give you feedback through pre- and post- tests. Uh, you did something that most beginners struggle with and that is aspirating the entire time which is very important. If you’re stabbing any needle you’re aspirating the entire time. So you did a great job of that. You did a great job of identifying the vessel. Excellent job, any questions?”

Once the coding scheme was set, the raters individually coded the transcripts achieving an interrater reliability (weighted Kappa) of 0.7.

### 2.1.3 Results

Based on the content analysis, the most frequently mentioned feedback provided during the pre- and post-test were, “Aspirating the entire time” which was referenced 17 times, “Inserting the needle at a 30-45 degree angle” which was referenced 14 times, and “Obtaining a clear image of the target vessels” which was referenced 12 times. A total reference count for all items is shown in Figure 2-1. Interestingly, ‘Hand position on the syringe’ and ‘Anchor hand on body to ensure stability’ were referenced 11 and 9 times, respectively, but are not part of the skills checklist. This may suggest the checklist is inadequate for providing feedback on all aspects of the CVC insertion procedure. A full definition for each item can be found in the Coding Handbook in Appendix D.
Based on this information, and the capabilities of the system, the authors focused future development of the learning system in two areas: (1) personalized learning feedback through the DHRT system, and (2) an introductory training video presented at the start of the DHRT system. The details of these items are presented in Phase 2.

2.2 Phase 2: Learning Interface Design and Development

The purpose of Phase 2 was to incorporate the information gathered in Phase 1 into a personalized learning interface for the DHRT system. Specifically, based on the results from Phase 1 of this chapter and the capabilities of the system, a personalized learning system was developed that involves 4 components: an instructional video, home screen with performance tracking, patient case information and needle insertion trial, and a feedback screen. The specifics of each of these components is provided in the remainder of this section.
2.2.1 Logging in and viewing the training video

The first step of the personalized learning system involves users logging into the system using a unique, de-identifiable code and watching a 7-minute-long training video created in Adobe Premiere. This video was developed to address the eight feedback items not addressed by the personalized learning interface, (see Figure 2-1). Specifically, the video provides an introduction to the system and highlights the skills on the CVC checklist that were identified in Phase 1 as being most frequently referenced. This includes a full demonstration of an IJ CVC placement on the DHRT, and demonstrations of how to hold the syringe, how to identify vessels on the ultrasound screen, and how to confirm vessel entry by aspiration. During these demonstrations, particular emphasis was placed on methods for correctly identifying the vessels, advancing the needle slowly and steadily, and selecting the appropriate site for venipuncture. After the demonstrations, the video explains each part of the learning feedback screen.

2.2.2 Home Screen and Performance Tracking

The ‘home screen’ (see Figure 2-2) shows a graph of an individual’s performance over time, provides them with general tips to improve their CVC insertion performance, provides navigation buttons that allow them to re-watch the training video or start simulation practice, and contains “sign out” and “close program” buttons. The graph showing increased performance was included to show users how much their performance has increased over their training because research indicates learning is enhanced when an individual perceives a skill as learnable rather than fixed.86
2.2.3 Simulation and Patient Profile Development

Once the ‘start simulation’ button is activated from the home screen, a hypothetical patient profile is displayed containing relevant clinical information in the form of 8 patient profiles, see Figure 2-3 for example screen. These profiles were developed by a second-year surgical resident and the chief of surgery at Hershey Medical Center and included characteristics of patients that would impact the ease of placing a central line. For example, Case 7 was,

“A 33-year old morbidly obese male with DM, COPD, and CHF presents with necrotizing fasciitis and requires a central line prior to operative debridement for hemodynamic support.”
The height and weight of the patients were developed to represent an appropriate BMI for each patient. For example, Case 7 was given the following anthropometry:

Height: 5’ 7”; Weight 282.2 lbs.” In this case, the large BMI, and therefore large amount of adipose tissue, would make placing a central line more difficult due to an increased depth of vessels. Heights and weights of participants were determined using data from NHANES\textsuperscript{116}. The variations in patient anatomy were represented in the simulated model through haptics and visuals presented to the user. Specifically, variations in skin thickness, which require differing amounts of force to puncture through the surface of the skin, were captured and presented to the user through the robotic arm of the DHRT system, see Pepley, et al.\textsuperscript{63} for details. A graph depicting the changes in force related to depth is shown in Figure 2-4.
In addition to changing the force profile, variations in adipose tissue (indicated by the patient’s weight) resulted in deeper or shallower vessels in relation to the surface of the skin which was apparent through the distance the needle was required to travel before puncturing a vessel as well as the visualization of the vessels on the ultrasound screen. In other words, the vessel depth, location, diameters and wall thickness were varied in each of these profiles based on realistic anatomical variations. The amount of force required to move through the skin, subcutaneous tissue, adipose tissue, and vessel walls was varied according to work by Gordan et al. The variation of vessel location for

Figure 2-4: Variations in Force for Profile 7: The amount of force required to insert a needle into the tissue will generally increase as the needle goes deeper in the tissue due to the friction forces between the needle and the tissue. How much the force increases depends on the type of tissue. For example, in this image there is a change in tissue type around 10cm and again around 15cm.
Profile 7 is shown in Figure 2-5 (see Appendix A for all variations) with the resulting ultrasound image.

After viewing the patient case information, users start the trial. A black screen will appear and they can begin the needle insertion procedure. The patient profiles are provided in order of easier to harder over the duration of the training. This was done because research has shown that skill transfer to novel situations is enhanced when variability of the scenario is included in skills training and learning is optimized when task difficulty is increased proportionate to the learner’s skill level.
2.2.4 Simulated Ultrasound Image

Once users press ‘start’ on the patient profile page and an ultrasound image appears, the user can lift the ultrasound probe and place it on the scanning surface. At this time, two vessels will appear on the screen as shown in Figure 1-4. These images will shift on the screen as the ultrasound probe is shifted to the left or the right of the scanning surface. Pressing down with the probe will cause the vessels to compress slightly. As the needle is inserted into the scanning surface, the needle can be tracked on the ultrasound image by noting where the horizontal lines across the screen are flexing as shown in Figure 2-6. When the tip of the needle enters a vessel and the user is aspirating (pulling back on the plunger of the syringe) a colored bar will appear across the top of the screen to indicate ‘flash’. A blue bar indicates the needle is in the vein (target) and a red bar indicates the needle is in the artery.

Figure 2-6: Virtual Ultrasound Image: The needle tip is visible in the right vessel. The blue bar across the top of the screen indicates that the needle tip is in the vein.
2.2.5 Personalized Learning Feedback Screen

Once the simulation is complete, a personalized screen is shown that captures the feedback from Phase 1 of the study (see Figure 2-7). The design of the personalized learning screen contains 7 unique boxes to address the items identified in Phase 1 as most commonly referenced. To address ‘conduct entire procedure without any mistakes’ (referenced 10 times) there is an overall grade in Box 1. This is calculated by a weighted grading system for each item on the CVC checklist. The weights for each item are based on the frequency of reference determined during Phase 1. Box 2 includes the overall difficulty of the scenario, as determined by a team of surgical experts. Box 3 includes the number of insertion attempts and alerts the user if they went through—and—though the vein. Although number of insertions was not analyzed as part of the content analysis, it was an item on the CVC checklist, so it was included as part of the feedback screen. An attempt was counted as a new insertion attempt if the user retracted the needle from the skin. Box 4 shows the average angle of insertion (referenced 14 times), which is calculated from the overall path created by the needle tip. Box 5 shows how closely the needle tip was to the center of the vein (referenced 6 times) based on the coordinates of the tip of the needle at the end of the trial relative to the center of the circular vessel. Box 6 indicates what percentage of time aspiration occurred and if there was any unnecessary movement. Finally, on the top right side of the screen is a high score board which keeps a history of top scores from Box 1. After the system was developed, a usability study was conducted to identify its effectiveness.
2.3 Phase 3: Usability Testing of the DHRT Personalized Learning System

The purpose of Phase 3 of the study was to evaluate the initial design of the learning interface through user testing and interviews. This was achieved through an experimental study with 8 surgical residents who were recruited from Hershey Medical Center. To qualify for the study, participants were required to have placed a central line in a clinical setting. They were remunerated with a $15 gift card. The details of the study follow.

Figure 2-7: Personalized Learning Screen for the Dynamic Haptic Robotic Trainer. Each box provides feedback on (1) the overall performance, (2) the difficulty of the patient case, (3) the number of insertions and if the needle passed through the back wall of the vein, (4) the angle of insertion, (5) the distance to the center of the vein, and (6) the frequency of aspiration.
2.3.1 Procedure

At the start of the study, the purposes and procedures were explained to the participants and informed consent was obtained. Next, participants were asked to perform trials on the DHRT system including signing in, watching a tutorial video, starting a simulation, performing a CVC insertion simulation, *verbally reporting their interpretation of the results* presented on the feedback screen, and then starting another case. It is important to note that the results presented to the user were not the results of any simulation they had completed, but were static for each user and created for the sole purposes of this study. After completing the study, participants completed a brief survey asking them to rate the following items on a scale of 1 to 5 with 1 being “least useful” and 5 being “most useful”: importance of competition, detailed feedback, graded feedback, feedback on aspiration, feedback on centering, feedback on angle, video training, and overall opinion. Finally, participants were allowed to provide suggestions on the system design.

2.3.2 Results

The results of the survey showed that all median survey responses for the 8 items on the usability survey were above a median value of 3 (neutral) with detailed feedback (M=5), feedback on centering (M=5), and feedback on angle (M=5) receiving the highest ratings (see Figure 2-8). The following suggestions were provided by participants during the study: (1) Present methods for improving parameters of performance in addition to scoring, (2) The use of grade letters and specific performance parameters is helpful, and (3) Provide a personalized tip for each user on which parameter(s) to improve based on their past performance.
The results indicated that the information currently being provided was useful and important, but that the format of the delivery of the information could be improved. Several modifications were made to the interface as shown in Table 2-1. Additionally, the placements of several buttons were changed to accommodate navigation errors made by users during testing. Because the original system was presented as a paper-prototype, the wording and button placements were changed during the study based on an as-needed basis. The ‘close program’ button was added to the learning interface so that users would not need to navigate back to the home page to close the program. The ‘try again’ button was removed for experimental design purposes when testing the effectiveness of the system as a whole (see Chapter 5 for details) to ensure that participants in the DHRT group and the manikin group where completing the same number of practice trials.

Figure 2-8: Median responses on feedback survey items
Table 2-1: The Changes of Each Box from Version 1 to Version 2

<table>
<thead>
<tr>
<th>Box #</th>
<th>Version 1</th>
<th>Version 2 (after user testing)</th>
<th>Reason for Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Grade 76%</td>
<td>Grade Score 99% 911.33</td>
<td>The grading system was changed from a maximum score of 100 to a maximum score of 1000 to capture more granular differences in performance.</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inserted into vein but need to improve angle of insertion</td>
<td>You have successfully accessed the jugular vein</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Difficulty:</td>
<td>Case Difficulty:</td>
<td>The case information was shortened to be more clear and direct.</td>
</tr>
<tr>
<td></td>
<td>⭐⭐⭐⭐⭐</td>
<td>- shallow vessels</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tip: For this case, you needed to insert the needle at a more shallow angle. However, in obese patients, a steeper angle may be required.</td>
<td>- large vessels</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- overlapping vessels</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>No changes were made.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Angle of insertion: 54°</td>
<td>Angle of insertion: 36°</td>
<td>The image was changed to provide more accurate visual feedback on the needle angle.</td>
</tr>
<tr>
<td></td>
<td>⭐⭐⭐⭐⭐</td>
<td>⭐⭐⭐⭐⭐</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Needle tip position: Near the center</td>
<td>Needle tip position: Perfectly centered</td>
<td>The image was changed to a dart-like image showing the exact position of the needle tip relative to the center of the vein.</td>
</tr>
<tr>
<td></td>
<td>⭐⭐⭐⭐⭐</td>
<td>⭐⭐⭐⭐⭐</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Unnecessary movement avoided:</td>
<td>You aspirated for 100% of the procedure</td>
<td>The feedback on aspiration was changed to show a percentage rather than a binary yes/no. Participants indicated that “unnecessary movement” was vague and confusing so feedback on this was removed.</td>
</tr>
<tr>
<td></td>
<td>⭐⭐⭐⭐⭐</td>
<td>⭐⭐⭐⭐⭐</td>
<td></td>
</tr>
</tbody>
</table>
2.4 Discussion

The purpose of this chapter was to discuss the methodology and results of a user study that was conducted in three phases to: (1) systematically analyze the verbal and written feedback given to individuals being tested on IJ CVC procedures (2) incorporate the information gathered in Phase 1 into a personalized learning interface for the DHRT system and (3) evaluate the initial design of the learning interface through user testing and interviews. This is important because there is a large gap in performance when transitioning from simulators to the first several surgeries on clinical patients which may be to the quality and type of feedback received during the training process. Therefore, DHRT Graphical User Interface (GUI) personalized learning system was designed to give consistent, appropriate, quality feedback for each practice trial. While practice is critical to learning a new motor skill, feedback, and the ability to incorporate that feedback into practice, is also a critical part of learning. The timing and type of feedback (task specific, processes specific, or self-regulatory) provided to an individual can change what and how they learn. The results of Phase 1 identified process-specific and task-specific feedback that was given verbally to participants. The items on which feedback was most frequently given were “Aspirating the entire time,” “Inserting the needle at a 30 to 45 degree angle,” and “Obtaining a clear image of the target vessels.” The learning interface was designed to present specific and detailed feedback on each item, after each trial in both images and text. Research suggests that information retention can be improved by presenting personalized feedback in both textual and graphical forms, giving corrective feedback about small errors rather than large ones, and giving feedback after good trials.

The results of the study identified a methodology for creating a learning interface based on verbal feedback for a surgical procedure. However, there were several limitations to this study. First, the feedback that was analyzed was given by a single second year resident. Although the individual giving feedback trained all incoming residents on CVC procedures there may still have been a bias in the content and consistency of feedback. Additionally, only 8 residents participated in the user study. Although residents conducted all training in CVC procedures at HMC, the content of the
feedback provided during the trials may have differed between individuals. Additionally, although participants were able to verbalize an understanding of the feedback being provided, no analysis of their actual performance or learning gains was conducted. Importantly, to overcome this limitation, a second study was conducted to determine if the DHRT system, including the personalized learning interface, is an effective method for training new medical residents in CVC placement (see Chapter 4). Future studies should seek to identify how these learning outcomes transfer to clinical settings.
Chapter 3
Comparing the Effectiveness of the DHRT System to Existing Manikin-Based Approaches


The previous chapter outlined the design and development of the DHRT personalized learning system. The goal of the current chapter is to provide support for the validation of the DHRT system as an effective method of training surgical residents on the skills needed to insert an Internal Jugular Central Venous Catheter (IJ CVC). This is important because the DHRT was developed to increase knowledge retention and increase effective skill transfer from simulators to clinical patients. The use of simulators reduces clinical errors but there is still a large gap in skill gains between training and performing the procedure on a clinical patient. This gap may be due in part to the static nature of training simulators; learning a motor skill on variations of a scenario results in better skill transfer than repeatedly practicing the skill in the exact same scenario. Additionally, many of the complications that occur during CVC procedures in clinical cases can be attributed to variations in anatomy. While Virtual Reality simulators like the DHRT system can present multiple realistic scenarios with variations in anatomy, which can potentially reduce clinical complications, they lack validation in surgical training and outcomes. Thus, the current chapter was developed to present a study aimed at identifying CVC insertion skill and self-efficacy gains between 3 different training groups (manikin training, DHRT training, half manikin and half DHRT training).

3.1 Methodology

The goal of this chapter was to determine if the DHRT system was an effective method of teaching CVC skills and if it increased surgical resident self-efficacy during CVC training. This is important because while virtual reality (VR) simulators have been
successfully used in other areas of surgical skills training\textsuperscript{58}, little research has been conducted on the use of VR for CVC skills training. Therefore, the current study was developed to answer the following research questions (RQ):

\textit{RQ1}. How, if at all, do medical students’ skill performance and CVC insertion self-efficacy changed after training with the DHRT, a manikin, or a combination of both training methods? It was hypothesized that students would improve their CVC insertions skills and self-efficacy when trained using any of the three methods as prior research has shown that simulator training increases resident comfort with CVC procedures\textsuperscript{20}.

\textit{RQ2}. Does the method of training impact gains in medical student CVC self-efficacy or CVC insertion skills? It was hypothesized that students who were trained using the DHRT or a combination of the DHRT and manikin trainer would have equal or better performance on CVC skills and self-efficacy over the course of the study because prior research has shown that VR training can improve trained skills\textsuperscript{28}.

\subsection{Participants}

A study was conducted with 18 third-year medical students between the ages of 23 and 35 (Mean = 26). Of these participants, there were 11 males and 7 females. Participants were recruited from the medical education program at Penn State Hershey Medical Center (HMC). Medical students who had prior experience with central line placement were ineligible for the study. Participants were randomly divided into groups of 3 (6 groups total) and assigned to one of three training conditions for each group. There were 17 right-handed participants and one left-handed participant. Participants were remunerated with a $15 gift card for the study, which lasted approximately 90 minutes.
3.1.2 Procedure

At the beginning of the study, the purpose and procedures were explained, any questions were answered, and informed consent was obtained. Participants were then asked to complete a Central-Line Self-Efficacy (CLSE) survey regarding their confidence in their abilities to perform central line insertion skills and any prior training they had on this or similar procedures (see Survey Instruments for details). Once complete, the group of three participants was given a demonstration of central-line placement from a second-year medical resident with expertise in the area using a Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) manikin. During the training session, participants were shown how to use an ultrasound, how to identify and distinguish between the artery and vein, how to use anatomical landmarks as a guidance for line placement, how to insert a needle, how to identify needle location based off of ultrasound feedback and how to confirm needle placement using flash feedback. Importantly, the CVC training procedures employed in the current study are the same procedures used to train individuals in the surgical residency program at HMC and the second-year resident who performed the training in the study had conducted numerous CVC training sessions at HMC.

Once the training was complete, each of the participants individually entered into a pre-test where they were asked to insert a needle for central line placement into the same Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) used for the demonstration while verbalizing what they were doing according to a think-aloud procedure. During the pre-test, each participant was observed and evaluated by the same second-year medical resident using a modified Internal Jugular Catheterization (IJ CVC) evaluation form (see Survey Instruments for details). Participants were not provided any feedback throughout their pre-test, but once the pre-test was complete, they were informed if they had successfully placed the needle and what errors occurred during the procedure. After receiving feedback, each of the participants individually participated in one of three training conditions as follows:

*Manikin training.* Participants in this condition performed all of their training on a second Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660).
Robotic training. Participants in this condition performed all of their training on the Dynamic Haptic Robotic Trainer (DHRT).

Mixed training. Participants in this condition performed half of their training (4 patient profiles) on the Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) and half of their training on the DHRT.

Before the training sessions began in the robotic and mixed training conditions, the participants were given a demonstration and a brief explanation on how the Dynamic Haptic Robotic Trainer (DHRT) worked and they were allowed to complete a practice trial on the device. This took about 10 minutes to complete. Next, in each of the conditions participants were provided with 1 of the original 8 patient profiles discussed in the introduction (See Patient Profile Development in Chapter 2).

Once the participant read the profile they were asked to place a needle in the center of the vein while using the think-aloud procedure. Unlike the pre-test, a researcher gave corrections as participants attempted the procedure. After the participant successfully entered the vein or accidently hit the artery they were provided with feedback on their performance (e.g. the final needle position, average insertion angle, number of attempts and insertion technique) and were able to ask any questions they had. Once complete, a new patient profile was provided and this procedure continued until all 8 profiles were completed. Each participant was given these profiles in a randomized order. This training took approximately 20 minutes.

When a participant had completed the 8 trials, they then completed a post-test assessment where they were asked to perform the CVC insertion on a Kyoto CVC Insertion Simulator II (Model # M93UB). This manikin was selected for the post-test because it differed in anatomical structure (e.g. depth and size of vessels, skin thickness, and vessel wall thickness) from both the pre-test and training manikin. The same procedures from the pre-test were used for the post-test and the same medical resident conducted this examination. After the test was complete, the participant was given extensive feedback on their performance and given the opportunity to see a demonstration of the rest of the central line placement procedure (e.g. insertion of the guide wire and catheter placement).

Finally, participants were asked to complete a post-training Central-Line Self-Efficacy (CLSE) survey regarding their confidence in their abilities to perform central line insertion skills after their training. In addition, participants were asked to complete a
Training Evaluation Survey (TES) regarding their experience with the training they received (see Survey Instruments for details).

### 3.1.3 Survey Instruments

Throughout the study, a series of three surveys were used to analyze participant performance and self-efficacy. The three surveys included a Central-Line Self-Efficacy Survey (CLSE), an Internal Jugular Catheterization Evaluation Form (IJ CVC), and a Training Evaluation Survey. The full versions of these surveys are included in Appendix B but details of are provided below.

*The Central-Line Self-Efficacy Survey (CLSE)* is a 14 question 5-point Likert scale survey dealing with participants’ confidence in their ability to perform the skills necessary to insert a needle for the central line procedure. The scale ranges from 1 (not at all confident) to 5 (extremely confident). Example items include, “Using tactile feedback during placement of the line”, “Modifying needle trajectory based on ultrasound feedback” and “placing the introducer needle at the center of the vein in one attempt”.

*The Internal Jugular Catheterization (IJ CVC)* evaluation form consists of 10 tasks including items such as “selecting the appropriate site for venipuncture”, “confirming vessel entry by aspiration of blood” and “conducting the entire procedure without any mistakes”. During the pre-and post-tests the tasks are marked as ‘pass’ (1) or ‘fail’ (0) by an independent evaluator. In addition, the number of attempts needed to insert the needle was also documented.

*The Training Evaluation Survey* is a 10-question 5 point Likert-Scale survey dealing with the participant’s perception of the training method they received. The scale ranges from 1 (completely disagree) to 5 (completely agree). These questions included items like, “The Haptic Robot (or manikin) Training … made me sensitive to patient anatomies’ impact on vessel location in the body” and “was an effective method for learning the CVC procedure”. The mixed training participants filled out surveys for both the manikin and robot.
3.2 Data Analysis and Results

In order to answer the research questions posed as part of this chapter, statistical analyses were performed on the pre- and post-test IJ CVC; pre- and post-training CLSE; and the training evaluation survey. These analyses and their results are presented in the following sections in relation to the research questions. All analyses were conducted using SPSS (v. 22.0) with an error rate of 0.05.

RQ1: How, if at all, do medical students’ skill performance and CVC insertion self-efficacy change after training?

The first research question was developed to understand how, if at all, each training method impacted CVC skill acquisition and participant self-confidence. In other words, did each method actually train participants in these skills. In order to answer this research question, Wilcoxon Signed Ranks test was conducted for each training method to compare participant responses on the 14-items from the pre- and post- CLSE surveys. The results revealed that for each of the training conditions, 13 of the 14 items were statistically different (p<0.05) indicating significant self-efficacy gains. For the manikin condition, there was not a significant learning gain for ‘using tactile feedback to help guide the introducer needle’ (z = -1.80, p=0.07). For the robot condition, there was not a significant learning gain for ‘Advancing and retracting the introducer needle slowly and steadily’ (z = -1.86, p = .06). Finally, for the mixed condition, there was not a significant learning gain for ‘using tactile feedback to identify the correct vessel for puncture’ (z = -1.91, p= 0.06).

In addition to comparing changes in the CLSE survey, an analysis was also conducted to understand if there were changes between the pre- and post-test IJ CVC. In order to do this, an exact McNemar test was computed for each of the 10 items on the IJ CVC. The results revealed that there were no significant differences between performances on the IJ CVC evaluation form for each of the training methods. Importantly, the average score for all participants on the pre- and post-test was 81.3% and 95.7%, respectively and all participants performed better on the post-test than on the pre-test. While not significant, it is interesting to note that the average post-test scores for the manikin, robotic and mixed training were 92.3%, 94.8%, and 100% respectively.
Additionally, it is interesting to note that the mixed training condition was the only condition with participants that did not make any mistakes on the post-test.

These results support the hypothesis that medical students would improve their skills when trained using any of the three methods and are promising for the use of the DHRT as a training method. None of the participants had ever been exposed to a haptic simulator yet still showed the same increase in learning gains as traditional training approaches. These results also support prior work which found that simulated training increases resident comfort with CVC procedures\textsuperscript{20} and support prior work that suggests virtual reality simulators are effective methods for surgical training\textsuperscript{28,31}.

**RQ2: Does the method of training impact gains in medical student CVC self-efficacy or CVC insertion skills?**

The second research question was developed to understand if different training methods impacted gains in self-confidence or acquisition of the skills necessary for the CVC procedure as judged by the IJ CVC. It was hypothesized that students who were trained using the DHRT or the mixed training method would have equal or better performance on skills and self-efficacy because prior research showed that VR training can improve trained skills\textsuperscript{28}.

In order to answer this research question, a MANOVA was computed with the independent variable being the difference between response for each of the 14 items on the pre- and post-training CLSE survey and the dependent variable being the training method (manikin, robot or mixed). The results of the MANOVA revealed that there were no significant differences in self-efficacy gains between the three training methods for each of the 14-items on the CLSE (\( p > 0.05 \)). In other words, no training method had a significantly larger (or smaller) impact on CVC self-confidence gains.

In addition to exploring the impact of the training method on student self-efficacy, the impact of the training method on performance gains between the pre- and post- IJ CVC evaluation was also explored. Specifically, a chi-square test of independence was performed to determine the relationship between these variables for each of the 10 items. The results showed no statistically significant results between CVC skill gains in each of the 10 items (\( \chi^2 (4, N=16) < 5.00; p > 0.144 \)). In order to compare the number of insertion attempts needed to complete the pre- and post-test IJ CVC, a repeated-measures
ANOVA was calculated. The results revealed no statistical significant differences between the three training conditions (F (1,2) =1.97, p < 0.184).

Finally, the training evaluation survey (TES) was analyzed to compare medical students’ feelings about the utility of the training methods for learning CVC procedures. Specifically, a Mann-Whitney U-test was computed with the dependent variables being the response to each of item on the TES and the independent variable being the method of training (manikin or robot only). Because participants in the mixed condition completed a training evaluation survey for both the robot and the manikin, the sample size for this analysis was 24. The results revealed a marginally significant difference between the conditions for the question “helped me understand how to modify CVC insertion procedures based on patient anatomy” (U=32.5, p < 0.058) with participants reporting a higher level of response to the robot than the manikin (Mean Rank was 14.05 and 8.95 for robot and manikin, respectively). There were no other significant findings. For individuals in the mixed condition, a Wilcoxon signed ranks test was computed to compare the responses to both training methods. The results revealed a significant effect for ‘helping me correctly identify when I had successfully inserted the needle’ (z=-2.06, P<0.04) with the manikin (Mean Rank = 4.6) being significantly higher than the Robot (Mean Rank = 3).

These results support the hypothesis that the DHRT or mixed condition would have equal self-efficacy after training and some support for higher self-efficacy and CVC performance gains in relation to modifying the procedure based off of patient anatomy. This supports prior research that shows virtual reality training can improve targeted skills. This is encouraging for the use of the DHRT as a training method, considering the pre- and post-test were performed using a manikin, possibly giving students in the robotic training method a disadvantage. However, the true impact of adapting to patient variability in a clinical setting remains to be seen.

3.3 Discussion

The goal of the current study was to understand how dynamic haptic training could be used to develop surgical skills. The main findings of this study are as follows: (1) Medical students had increased confidence (self-efficacy) in their ability to perform the skills necessary to insert a needle for the central line procedure when using the DHRT.
system; (2) There was no difference in learning gains or confidence gains between training methods, and; (3) Medical students reported that the DHRT training helped them understand how to modify CVC insertion procedures based on patient anatomy more than the Manikin training. This research suggests that students can be trained on CVC skills using haptic based training just as effectively as static simulators. An advantage of using haptic simulators is they can present an unlimited number of patients. This exposure to anatomical variation may reduce complications or improve new resident performance. It is important to note that the DHRT is not designed to replace the manikin training system which may still be necessary to learn other components of the CVC procedure such as sterile draping, but rather to enhance users’ ability to detect and adapt appropriately to anatomical variations during needle placement.
Chapter 4
Can Haptic Simulators Distinguish Expert Performance? A Case Study in Central Venous Catheterization in Medical Education


While Chapter 3 highlighted the ability of the DHRT system to increase self-efficacy and CVC skill gains in novice users, it did not identify if the DHRT system impacted CVC expertise development. The first step in identifying if they system impacts this expertise development is determining what characteristics of expert performance the DHRT system captures. Once this is attained, it is then possible to use these measures to identify how training improves novice skillsets in these areas. Importantly, the metrics gathered on the DHRT system are objective in nature in that they focus on items such as time to complete and average angle of insertion which are automatically computed in the system. This is important because current evaluation metrics are typically subjective in nature and involve a supervisor observing the skills and filling out a binary yes/no checklist to indicate if each step of a procedure was completed. This type of feedback system is useful for determining if participants know the appropriate steps but it may not be effective for evaluating their current skill level in each of these steps. In addition, this type of feedback many not be detailed enough to support the development of expertise. The type of feedback provided during the learning process influences what is learned which means that finding ways to provide detailed and accurate, objective feedback is crucial, and this may be possible through the use of motion tracking.

While not yet widely applied to Central Venous Catheterization training, there has been some development of tools for using motion tracking to quantify performance such as the Imperial College Surgical Assessment Device (ICSAD). The use of motion analysis has been used to identify objective metrics for comparing novice and expert performance in multiple medical fields. For example, Howells et al. looked at path length in Arthroscopy and found that experts had a significantly shorter path length than novices. Kim et al. looked at motion analysis specifically in CVC placement and Varas et al. validated motion tracking as means of assessing movement in manikin
simulated CVC placement\textsuperscript{104}. Both of these studies found that experts had significantly less hand movement than novices. Another study validated the use of electromagnetic motion tracking to distinguish between expert and novice performance by looking at path length, translational movement, and rotational movement of the needle hand as well as movement of the ultrasound hand\textsuperscript{107} and found that experts had significantly less movement. While the research across multiple fields have found that the overall path length and hand movement is less for novices, research has also indicated that experts may have a wider path of movement and more controlled rapid motions\textsuperscript{108}. However, this type of analysis has yet to be applied and validated for CVC placement in the DHRT system and thus it is not known whether or not this type of motion analysis can distinguish expert and novice performance on these high-fidelity simulators.

### 4.1 Methodology

The goal of this chapter was to validate the use of the Dynamic Haptic Robotic trainer (DHRT) for training CVC skills by comparing the performance of experts and novices. This is important because prior research has shown that experts will outperform novices when completing a procedure on a simulator\textsuperscript{110,111,120}. Therefore, the current study was developed to answer the following research questions (RQ):

**RQ1.** Does the experience level of the participant (expert or novice) affect individual performance on the path length of the needle tip, the standard deviation of the deviations from an ideal path, the proximity of the needle tip to the center of the vein, the velocity of the needle tip, or the time required to complete the procedure when using the DHRT system during CVC training and controlling for patient case (one of the 8 anatomical variations)? It was hypothesized that experts would have a shorter path length\textsuperscript{102,105}, smoother motions\textsuperscript{103}, closer proximity of the needle tip to the center of the vein, higher velocity of the needle tip\textsuperscript{108} and shorter time required to complete the procedure\textsuperscript{104} as supported by prior literature.

**RQ2.** Can expertise be distinguished during CVC placement using the DHRT system through motion analysis? It was hypothesized that expert performance on the DHRT system would be distinguishable from novices based on the path length of the needle tip, smoother motions, the proximity of the needle tip to the center of the vein, the
velocity of the needle tip, and the time required to complete the procedure. This hypothesis was developed because prior studies have looked at path length, hand stability, velocity and time to complete as measures of expertise\textsuperscript{103,104,107}.

**RQ3.** Do novices improve their CVC insertion performance through use of the DHRT system? It was hypothesized that novices would have a learning curve associated with improved performance when training on the DHRT system because prior work has identified learning curves for novices during CVC placement\textsuperscript{19}. It was also hypothesized there would be no learning curve for experts on the DHRT due to the fact that they are experts in the procedure; thus, any learning curves present for expert users would be attributed to learning of the system (DHRT) rather than learning of the procedure.

### 4.1.1 Participants

The participant data for the current chapter utilized the data for the portion of medical students who trained on the DHRT system in Chapter 3 including 8 male and 4 female novice participants between the ages of 23 and 35 (Mean age of 26). In addition, a second round of data collection was also conducted with experts recruited from HMC including 2 males and 2 females. To qualify expertise, all ‘expert’ participants had placed more than 100 central lines in their careers. The remainder of this section describes the procedure the expert users went through as part of the current investigation. The procedures for the novice participants can be found in the *Procedure section of Chapter 3*.

### 4.1.2 Procedure

The experts preformed similar procedures to the novices outlined in study 1 with a few minor modifications. Like the novices, at the start of the study, the purposes and procedures were explained to the participants and informed consent was obtained. Next, each expert was individually given a demonstration and a brief explanation on how the DHRT system worked. Then they completed a practice trial on the device, similar to the
novice training. They completed, in the same order, each of the 8 profiles presented to the novices. There was no pre- or post-test associated with the experts.

4.1.3 Metrics

In order to evaluate CVC skills during each trial, the following metrics were developed based on prior literature that looked at the position and movement of the tip of the needle of the DHRT system during each trial. Specifically, time to complete the insertion procedure, average velocity of the needle, distance to the center of the vein at the end of the trial, total path length of the needle during the insertion attempt, and the standard deviation of the deviations of the needle from an ideal insertion path were calculated. This section serves to highlight how each of these measures were calculated and provide rationale for their use in this investigation.

Average velocity of the needle tip was chosen as a metric because prior research has indicated that the velocity of hand movements can distinguish between expert and novice performance\textsuperscript{107,108}. However, velocity of the needle tip has not been studied to date in CVC procedures. Average velocity was defined as the average velocity of the needle tip when it was below the skin surface. This was calculated based on time required to travel the distance between each set of points recorded by the DHRT system and divided by that distance.

Distance to the center of the vein was developed as a metric because ‘placing the needle at the center of the vein’ is a component of the IJ CVC checklist and was identified as a metric receiving frequent feedback during Phase 1 of the study in Chapter 2. However, it has not been studied in prior research, likely because there is not currently an objective means of assessing performance on this skill while using a manikin simulator. Distance to the center of the vein was calculated as the radial distance from the tip of the needle at its final position to the center of the vein.

Needle tip path length was chosen as a metric because prior research has indicated that overall path length of hand motions can distinguish between expert and novice performance\textsuperscript{102,105}. However, path length of the needle tip in CVC insertion procedures has not been studied to date. Thus, path length was calculated as a summation of movement in the X, Y and Z coordinate system based on the points recorded by the DHRT system using the Euclidian distance formula. For the purposes of the current
study, path length began when the “needle tip” of the DHRT system was two centimeters above the “surface of the skin” because movement above that range was typically unrelated to the insertion procedure (e.g. picking up the syringe from its holster). See Figure 3-1 for a sample of a 3-dimensional plot of path length for both an expert and a novice.

Finally, Standard Deviation of the Deviations was calculated in an effort to measure the smoothness of the needles motion. This was calculated because prior literature has shown that experts have smoother motions during surgical procedures compared to novices\textsuperscript{103}. This metric was quantified by calculating the standard deviation of the deviations from an ideal path for a given vessel depth and average angle. The ideal path was defined as the path the tip of the needle would travel if moving in a perfectly straight line from the point of entry at the “skin surface” to the final resting point in the vessel. Ideal path was calculated based on the vessel depth and the average angle of

Figure 3-1: Needle paths of both an expert (blue) and a novice (red). The surface of the skin is a flat plane at $Y = 0$ (not pictured here). The blue tube represents the location of the vein and the red tube represents the location of the artery. The expert inserted slowly and then made micro-adjustments at the end of the insertion to precisely place the needle. The novice inserted the needle and missed the vein because they were lateral (too far in the positive x-direction), retracted the needle all the way out, and then inserted the needle again.
insertion, (see Figure 3-2). In the first image, the path length is shown as an equivalent to the hypotenuse of a triangle where $\theta$ is the average angle of insertion and the height of the triangle, $H$, is the vertical distance between the surface of the skin and the vertical depth of the final position of the tip of the needle. An ideal path length, or hypotenuse, was then derived using trigonometry (see Figure 3-2).

![Figure 3-2](image)

Figure 3-2: (A) The triangle derived from needle path, insertion angle $\theta$ and vessel depth $H$. (B) Two needle paths with similar standard deviations but dissimilar standard deviation of the deviations.

### 4.2 Data Analysis and Results

The metrics were analyzed using Statistical Package for the Social Sciences software, version 24 (SPSS, Chicago Ill) with an error rate of 0.05. Specifically, a MANOVA was used to determine if there was an interaction effect between experience level and patient case and to determine if either experience or anatomical configurations had an effect on performance using two independent variables (experience level and case number) and five dependent variables (path length, standard deviation of the deviations, distance to the center of the vein, velocity, and time to complete). In addition, a regression analysis was used to identify which of the five-metrics distinguished expert and novice performance. Finally, curve estimates were used to determine if there were
learning curves present for the five metrics across expertise level. The following section outlines the findings with relation to the research questions.

**RQ1.** Does the experience level of the participant (expert or novice) affect individual performance on the path length of the needle tip, the standard deviation of the deviations from an ideal path, the proximity of the needle tip to the center of the vein, the velocity of the needle tip, or the time required to complete the procedure when using the DHRT system during CVC training and controlling for patient case (one of the 8 anatomical variations)?

The purpose of the first research question was to understand if expert performance was significantly different from novice performance on any one, single metric. In order to answer this research question, a MANOVA was conducted and the mean and standard deviations of the 12 novices and 4 experts were computed for each of the five metrics (means and standard deviations can be viewed in Table 3-1). Overall, the MANOVA results revealed that experience had a significant main effect on the combined dependent variables $F(5, 78) = 4.393$, $p = .001$, Wilks' $\Lambda = .780$, partial $\eta^2 = .220$. In addition, there was a statistically significant main effect of expertise for time to complete $F(1, 3496) = 5.585$, $p = .020$, partial $\eta^2 = 0.64$, observed power of 0.646. Specifically, experts ($54 \pm 23.83$) performed the trials significantly faster than novices ($66.35 \pm 27.07$). All other variables were not statistically significant for expertise ($p > .1$). There was also no interaction effect between experience and patient case on the combined dependent variables, $F(35, 410) = 1.228$, $p = .182$, Wilks' $\Lambda = .598$, partial $\eta^2 = .098$. Importantly, assumptions were checked prior to using the MANOVA which revealed that the variables did not have multicollinearity (Tolerance $> .3$ and VIF $< 3.2$), which occurs when two or more independent variables are highly correlated. In addition, the independent variables were found to be linearly related to the logit of the dependent variable according to the Box-Tidwell (1962) procedure and a modified significance of 0.0045. The data was examined for outliers and 3 significant outliers were removed prior to analysis.
Can expertise be distinguished during CVC placement using the DHRT system through motion analysis?

While the previous research question investigated how expertise affected performance on the metrics, the purpose of this research question was to understand if any combination of any or all of the metrics can be used to distinguish between expert and novice performance. In other words, can performance on the metrics be used to predict expertise. In order to answer this question, a binomial logistic regression model was computed in order to determine if the path length, average velocity, time to complete, deviation of the deviations, and distance to the center of the vessel, were predictors of expertise. The model was found to be statistical significant $\chi^2(5) = 72.477$, $p < .0001$ with the model explaining 78.5% (Nagelkerke $R^2$) of the variance and correctly classified 91.7% of the cases. Specifically, the model showed that expertise could be predicted by path length ($p < 0.001$), average velocity ($p < 0.002$), time to complete ($p < 0.001$), and the standard deviation of the deviations ($p < 0.031$), see Table 3-1 for means and standard deviations. Distance to the center of the vessel was non-significant ($p > 0.614$). In other words, this model showed that while there isn’t a single variable that predicts expertise, these four variables are significant contributors to determining

<table>
<thead>
<tr>
<th></th>
<th>Novice</th>
<th>Expert</th>
<th>MANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Path Length</td>
<td>88.274 cm</td>
<td>56.63</td>
<td>93.68 cm</td>
</tr>
<tr>
<td>Standard Deviation of the Deviations</td>
<td>0.1946</td>
<td>0.1222</td>
<td>0.2438</td>
</tr>
<tr>
<td>Average Velocity</td>
<td>2.58 m/s</td>
<td>0.3105</td>
<td>2.56 m/s</td>
</tr>
<tr>
<td>Distance to the Center of the Vessel</td>
<td>.6333 cm</td>
<td>0.3667</td>
<td>.5606 cm</td>
</tr>
<tr>
<td>Time to Complete</td>
<td>66.35 sec</td>
<td>27.07</td>
<td>54.24 sec</td>
</tr>
</tbody>
</table>

Highlighted values indicate significance at $p<.05$
expertise on the system. Importantly, the sensitivity of the model was 79.2%, the specificity was 95.8%, the positive predictive value was 86.3%, and the negative predictive value was 93.2%. Assumptions were checked prior to computing the model. Specifically, five outliers were removed from the sample due to having studentized residuals values greater than 5 standard deviations away from the mean.
The final research question was developed to understand if a learning curve was present for novices in each of the metrics when using the DHRT system to improve their CVC performance. In order to answer this question, a curve estimate was fitted for each of the four significant factors from the regression analysis in RQ2. Specifically, a curve-fit estimation was performed to determine if there was a statistically significant change in performance over time by plotting the four metrics (path length, average velocity, standard deviations of the deviations and time to complete) separately for novices and experts against the order in which each case was presented, see Figure 3-3 for example plot. Next, linear, logarithmic and power curves were fit to each data set and separately analyzed for novices and experts to determine significance. Importantly, there were no significant findings for the experts. This indicates there was no learning of CVC skills for experts on the DHRT system, as was expected. However, for novices, there was a weak
but significant negative linear model for Path Length F(1,70) = 14.274, p < .001, R^2 = .169; and a moderate, significant power model for Time to Complete F(1,70) = 39.482, p < .001, R^2 = .36. This indicates that novices improved these specific skills when using the DHRT system over 8 trials.

Figure 3-3: Change in path length over time. Experts (top) showed no significant decrease or increase while novices (bottom) had a significant linear model for decrease in path length over time.
4.3 Discussion

The goal of the current chapter was to understand and compare expert and novice performance on the DHRT and identify objective metrics for evaluating CVC performance. This is important because research indicates that, even after training, there is still a large skill gap between expert and novice performance. The results of the empirical study indicated that experts were significantly faster at inserting the CVC needle in the DHRT system. In addition, needle path length, standard deviations of the deviations of the needle path, time to complete the needle insertion and average velocity of the needle together were able to predict expertise in the system. These findings suggest that the combination of these four metrics can assess skill acquisition on CVC simulators allowing for the development of systems that direct novices to engage in more deliberate practice of these skills, which is a critical part of learning.

It is important to note, however, that the findings on path length contradict prior research which suggested that experts have a shorter needle path length and fewer needle movements during CVC insertions. However, this prior work looked at the movement of the entire hand over the course of the procedure on a manikin rather than the movement of the needle inside the ‘tissue’. On the contrary, path length in the current study was calculated starting 2cm above the ‘skin’. This is significant because the experts in the current study deployed a variety of strategies for locating the needle tip when inside the tissue. For example, one expert would slightly jiggle the needle up and down as they inserted it into the skin and through the tissue in attempt to carefully and methodically identify and follow the needle tip on the ultrasound image. This type of technique would cause increases in both path length and needle movements (standard deviations of the deviations) under the surface of the skin. It is also interesting to note that these findings contradict prior work in Arthroscopy that found that experts had a significantly shorter path length and were significantly faster than novices. These findings highlight the importance of validating the use of needle path length for specific applications before systematic application.

The results also showed that novices had significant improvements in their path length and time to complete through the 8 trials completed on the DHRT system as demonstrated through the curve analysis. This is promising because it shows that they are improving these skills through the use of the system, even when presented with a limited number of cases. This finding supports prior survey-based research that found the DHRT
system to be an effective method of training medical students\textsuperscript{121} and that the DHRT system can improve novice insertion performance\textsuperscript{63}. In addition, there were no improvements on the four-examined metrics for experts who used the system. This suggests that experts are able to apply their CVC insertion expertise to complete trials in the DHRT system with relative ease. These findings are in line with prior research in dentistry\textsuperscript{120} and laparoscopic surgery\textsuperscript{110} that showed that experts outperformed novices on haptic feedback systems. These findings suggest that the DHRT system can be used as a method for improving key aspects of CVC expertise development and support the use of high-tech simulators in CVC training.

There were several limitations to this study. Specifically, there was a limited number of novice and expert participants and a limited number of trials completed which may limit the generalizability of the findings. Because of this, future research should be geared at larger studies that not only look at short-term skill gains, but long-term implications in the surgical suite. Second, the study here only looked at skill gains and expertise measurements on the DHRT system and not on the transfer of these skills to the operating room. This is an important area of future work as prior research has shown that even the most realistic models only temporarily improve CVC procedural skills\textsuperscript{21} and this skill acquisition declines over time\textsuperscript{21}. Therefore, future work should be geared at looking at the skill transfer of the DHRT system to the surgical suite.

Encouragingly, experts were able to perform at a constant high level of performance throughout the DHRT system indicating that it may be picking up on some of the key skills needed to be successful in the operating suite. The use of medical students may also limit the results of the study as they have no prior training in surgical skills. However, it is important to note that the medical students went through similar training procedures presented to first year residents in the ‘resident bootcamp’ presented at Hershey Medical Center. Finally, the study was limited by the number of practice attempts completed and did not look at the optimal number of trials needed to reach expert performance or plateaus in the learning curve. Thus, future work should be directed not only at optimizing the design and development of CVC high-tech simulators, but also the learning programs that surround their development.
Chapter 5
An Exploration into the Impact of Simulator Fidelity and Feedback on Surgical Resident Subjective Performance in Central Venous Catheterization Training

Publication: Mary Yovanoff, David Pepley, Katelin Mirkin, Jason Moore, David Han and Scarlett Miller. “An Exploration into the Impact of Simulator Fidelity and Feedback on Surgical Resident Subjective Performance in Central Venous Catheterization Training” To be Submitted to The Journal of Academic Medicine, July 2017.

While the previous chapters highlighted the development of the DHRT personalized learning interface and the effects of variations in short (30 minute) CVC training sessions (manikin versus robotic) on self-efficacy gains and expertise development of medical students, the current chapter turns the focus to the impact of long-term CVC training (6 months) on surgical resident CVC self-efficacy gains and self-ratings of performance. This is because both the results from Chapter 3 and prior research have suggested that learning on a simulator can increase resident comfort with CVC procedures. Self-efficacy is an important construct to explore in surgical residency education because there are links between positive self-efficacy and increased skill performance. Specifically, this prior work has shown that, when learning a new motor skill, trainees who have increased self-efficacy also have increased motor skill acquisition; in other words providing trainees with either positive feedback such as ‘this was much better than last time’ (to increase self-efficacy) or negative feedback such as ‘this is still wrong’ (to decrease self-efficacy). In addition, individuals who believe that they can adequately and effectively perform a skill will be more likely to do so regardless of actual past performance.

Research has also suggested that self-efficacy may not be a good predictor of future performance. Specifically, this research has shown that the relationship between self-efficacy and performance decreases when applying the task to novel situations. This is important because surgical residents must apply the skills and knowledge they gained while practicing on a simulator to the novel situation of performing the procedure on an actual patient. However, the current low-fidelity simulators used in CVC education only provide a single anatomical patient configuration for surgeons to practice on. While this allows residents to ‘simulate many’ CVC needle sticks without consequence, they are
simulating the *exact same* patient scenario. While this type of practice is important for initially learning a skill such as CVC needle insertion, repeatedly practicing the same skill in the same scenario may not increase skills pass a basic proficiency level\textsuperscript{51}. Because of this, practicing the procedure on actual patients after learning on a low-fidelity CVC simulator can be seen as a novel surgical scenario. Additionally, when high performance is expected, as is the case in the surgical suite, skill performance may break down despite high self-efficacy or a resident may “choke under pressure”\textsuperscript{57}.

In addition to exploring self-efficacy gains, self-ratings of performance are also important to consider because previous research has shown that motor skills and self-perception of those skills *improve* with training\textsuperscript{82,87}, but novices have a low ability to judge skill performance\textsuperscript{55,56}. Therefore, while self-perception may improve with experience, and individuals with higher objective scores of performance are better at self-assessing their performance, a literature review of 17 studies on physician self-assessment of skills found that there is little or no relationship between self-assessment and objective measures of performance\textsuperscript{97}. More specifically, research conducted on training novices on a knee joint injection, found that self-ratings were directly related to confidence, but were *inverse*ly related to objective measures of performance\textsuperscript{98}. In other words, participants were overconfident in their ability to perform the task and were also inaccurate in assessing their own performance on the skill\textsuperscript{98}. This may be due, in part, to the fact that when learning a relatively simple motor skill, brain scans have revealed that learners may be cognitively lazy and complete several repetitions of the skill with only minimal cognitive effort. Specifically, these scans revealed that participants did not engage the portions of their brains which are utilized for memory consolidation and, therefore, skill retention\textsuperscript{90,91}. Therefore, if residents are confident in their ability to perform the task and are rating their performance high, they may practice performing the procedure without cognitively engaging in the process and therefore will not learn or retain the skills needed to perform the procedure in a clinical situation. This ‘laziness’ can be combated by giving novices progressively harder situations to practice on\textsuperscript{52} and by providing accurate, detailed feedback on performance\textsuperscript{75}.

One advantage of using a VR (virtual reality) system over a standard manikin simulator is that they can be designed to present personalized feedback in both textual and graphical forms which can improve information retention\textsuperscript{79}. Additionally, well-designed VR systems can address all three necessary stages of surgical skill acquisition: cognitive (knowledge), associative (technical skill), and autonomous (adequate
judgement) skills by providing consistent objective feedback. Providing feedback after good or accurate performance increases both skill performance and self-efficacy separately and additional research shows that increased self-efficacy directly impacts skill performance and motor skill acquisition.

5.1 Methodology

The DHRT system was designed to expose residents (novice trainees) to diverse training scenarios and provide consistent, objective feedback through a personalized learning interface. In order to validate its use as a learning tool in CVC procedures, the goal of the current chapter was to determine how self-efficacy and self-ratings of performance were effected by training on either the DHRT or manikin system. Specifically, the current study was developed to answer the following research questions (RQ):

RQ1. Do surgical residents improve their CVC self-efficacy over the course of training and is this improvement dependent on the type of training provided? Specifically, this research question sought to understand if resident CVC self-efficacy changed from pre- to post- CVC training and if this difference was due to variations in the CVC training method (manikin or robotic). It was hypothesized that self-efficacy would increase throughout training because prior work has shown that simulator training increases resident comfort with CVC procedures. However, it was also hypothesized that residents in the robotic group would have equal or better improvements on CVC self-efficacy over the course of the study because prior research has shown that providing feedback after good or accurate performance increases self-efficacy, and providing learners with augmented feedback (or feedback not normally present) can help engage cognition during the learning process. The DHRT was designed to provide detailed specific, and objective feedback which has been shown to improve skill gains and retention while feedback on the manikin remained subjective and delivered at the discretion of the observer.

RQ2. How do perceived ratings of performance change throughout training and is this dependent on the training method? Specifically, this question sought to address how
participant self-ratings of performance compared across training groups over the duration of training. It was hypothesized that self-ratings of performance would increase throughout training because prior studies have indicated that self-perception of motor skills improve with training\textsuperscript{82,87}. It was hypothesized that the robotic training group would have lower self-ratings of performance compared to the manikin group because the robotic group was provided with consistent objective feedback on performance. This was hypothesized because prior research has indicated that novices have a low ability to judge their own skill performance and are often over-confident in their abilities\textsuperscript{55,56}, but when provided with realistic feedback they may be better equipped to accurately assess their own performance\textsuperscript{96}.

\textbf{RQ3.} How do perceived ratings of performance on the robotic trainer relate to objectively calculated performance scores? Specifically, this question sought to address how participant self-ratings of performance compared with the objective score generated by the DHRT system. While previous research has indicated that even experts may have a low ability or no ability to judge their own skill performance\textsuperscript{55,56,97}, providing individuals with feedback during training may improve their ability to accurately assess their own performance\textsuperscript{96}. Thus, it was hypothesized that there would be a weak positive correlation between self-ratings of trial performance and scores generated by the DHRT system.

\textbf{RQ4.} How do perceptions of the effectiveness of the training simulator differ between simulator fidelity? Specifically, this question sought to identify if there were any differences between training groups when participants were asked to evaluate how effective the training they received was for increasing the skills required to complete a CVC procedure. It was hypothesized that participants in the robotic training group would rate their training method as equally effective or more effective than participants in the manikin training group because prior research has shown that VR training systems can be effective for teaching needle insertion procedures\textsuperscript{31}. Additionally, the results presented in Chapter 3 demonstrate that individuals trained on the DHRT system rate it to be just as effective as those trained on the manikin.
5.1.1 Participants

Participants were recruited from the 1st year residency program at Hershey Medical Center (HMC). Participants were invited to participate in the study during their “residency boot camp” which occurs during the first week of their surgical residency. There was a total of 26 participants (6 female, 20 male). The specialties of the residents were as follows: General Surgery (6), Preliminary Residents (9), Orthopedics (5), Otolaryngologists (2), Urology (2) Plastic Surgery (2) and Vascular Surgery (1). There was one left-handed participant. It should be noted that three participants (all male) were unable to finish the study due to constraints on their residency training. Of the three participants, one was in the robotic training group and two were in the manikin training group. Their data was included in the analysis when available. Data was not available for the post-test IJ CVC skills test, Post-test motion tracking, post-test Central Line Self Efficacy Survey, or the Training Evaluation Survey (see the following section for a description of each of these metrics). Additionally, one participant from the manikin training group only completed 12 out of the 22 required needle sticks.

5.1.2 Procedures

The study was conducted as part of the surgical residency program at Hershey Medical Center. During the first day of boot camp, the research team provided a brief oral summary of the proposed research and participants were informed that they would receive the same CVC training and testing regardless of their participation in the study, but if they chose not to participate, their data would not be collected or analyzed as part of the study, and the outcome of their residency would not be impacted in any way. Informed consent was then obtained and all questions were answered. Participants were then asked to complete a Central-Line Self-Efficacy Survey (see Appendix A) regarding their confidence in their abilities to perform central line insertion skills and any prior training they had on this or similar procedures (See Survey Instruments in Chapter 3 for details). As part of their general training, participants then viewed an 18 minute video on central line placement. Beyond watching the video, participants were given no initial instructions on how to perform the procedure.
Once the video had been viewed, each of the participants individually entered into a pre-test where they were asked to insert a needle for central line placement into a Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) while verbalizing what they were doing according to the think-aloud procedure\textsuperscript{114}. During the pre-test, each participant was observed and evaluated by a second-year medical resident using a modified Internal Jugular Catheterization (IJ CVC) evaluation form. The IJ CVC evaluation form is a 10-item checklist focusing exclusively on the needle insertion portion of the procedure and including items like “continuously aspirating the entire time”, “conducting the entire procedure without any mistakes”, and ”selecting the appropriate site for venipuncture” (see Appendix B for full checklist). Participants were not provided any feedback throughout their pre-test, but once the pre-test was complete, they were informed if they had successfully placed the needle and what errors occurred during the procedure. There was a 5-minute time limit for the pre-test and if the resident did not successfully stick a vessel during the 5-minute time period, the pre-test was terminated. Importantly, the needle pathway was tracked using an Ascension 3D Guidance trakSTAR (Sheldburne, VT) electromagnetic position tracking system.

Five weeks after the pre-test, participants were given a demonstration of the full IJ CVC procedure and then were asked to performed 2 practice insertions on which they were given feedback. Specifically, participants were given a demonstration of central-line placement from a second-year medical resident with expertise in the area using the Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) manikin. During the training session, participants were shown how to use an ultrasound, how to identify and distinguish between the artery and vein, how to use anatomical landmarks for line placement, how to insert a needle, how to identify needle location based off of ultrasound feedback and how to confirm needle placement using flash feedback. Importantly, the CVC training procedures employed in the current study are the same procedures used to train individuals in the surgical residency program at HMC and the second-year resident who performed the training in the study had conducted numerous CVC training sessions at HMC.

After the training session, each of the participants was randomly assigned to one of two training conditions: manikin or robotic.
Manikin training (N=13). Participants in this condition performed all of their training on a Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) or a similar training manikin.

Robotic training (N=13). Participants in this condition performed all of their training on the DHRT. Importantly, all of the participants assigned to the robotic group were right handed due to current limitations in the DHRT system.

Participants followed the same procedures regardless of their assigned training group. Specifically, after viewing the demonstration, participants were asked to individually complete 2 practice insertions on their assigned training devices. There were 2 DHRT trainers available and 1 manikin trainer available for this session, all of which were in separate rooms. Participants in both groups were provided with a Self-Evaluation Training book (see Appendix D) to complete throughout their training program. For the first two needle sticks, participants were asked to individually perform the needle insertion using an ultrasound and a Blue-Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660) manikin. A second-year surgical resident observed their performance and provided verbal feedback. After receiving feedback, participants were asked to complete a Practice Insertion Performance Evaluation Form (see Figure 5-1) from their Self-Evaluation training book which asked them to rate their performance, the number of attempts it took them to access the vein, any errors they made, and if they received feedback. The full book can be found in Appendix E. Once they completed the two sticks, the initial training session was complete. Importantly, the number of manikin trainers was limited by the availability of a qualified surgical resident to oversee the practice and provide verbal feedback on performance.
Answer the following questions before continuing to the next trial:

<table>
<thead>
<tr>
<th>Rate your performance</th>
<th>Very Poor</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Great</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of attempts* to access vessel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

List any errors you made:

Did anyone give you feedback? □ YES □ NO

* Attempts are defined as a large manipulation with the needle or withdrawing the needle and re-entering

Figure 5-1: Practice Insertion Performance Evaluation Form included in Self-Evaluation Training Books

Prior to the first two practice needle sticks for the robotic group, participants in this condition were shown an instructional video on how to use the DHRT system (see Logging in and viewing the training video in Chapter 2), as the DHRT system was designed so that an individual with no prior instruction could use it without outside help. Next, participants were asked individually to complete their first two needle sticks on the DHRT system after creating their user profile. Prior to each of the two needle insertion practices, participants were presented with a baseline patient profile (see Chapter 2 for details). After reading the patient profile, participants were presented with a virtual ultrasound screen and were asked to complete the needle insertion procedure using the DHRT system until they were confident in the placement of their needle at which point they could end the trial by placing the ultrasound probe in a red box as shown in Figure 5-2.
Once the participant ended the trial, a screen prompted them to provide a performance rating on a scale of 1-5 (1 being ‘very poor’ and 5 being ‘very good’) as shown in Figure 5-3. After providing their rating, participants were provided with a feedback screen as shown in Figure 2-11 (see Chapter 2). Next, like the manikin group, the robotic group also recorded their feedback and performance on the Practice Insertion Performance Evaluation Form. After viewing the feedback screen and completing this assessment, participants were able to complete their second trial. While a researcher was present in the DHRT rooms in case of glitches or problems operating the system, they provided no feedback to the participants on CVC performance.
About 4 weeks after the two initial sticks were complete, a second training session was conducted in which participants were instructed to complete 10 needle insertions. They were given no additional instructions, but they did receive feedback either orally from a resident in the manikin group or through feedback by the DHRT system just as they did during the two initial sticks. After each of the 10 completed trials, participants filled out the same Practice Insertion Performance Evaluation form used in the first two-sticks. For the robotic group, the first and last trial were the baseline patient case and the middle 8 trials presented variations in anatomy and patient demographics and increased in difficulty over time. The manikin group received no such case data. It is important to note that 11 participants were not able to complete their initial 10 practice sticks during this time frame because they were on vacation, were

Figure 5-3: The Performance-rating Screen for the DHRT was presented to participants immediately after completing a trial but prior to receiving any feedback on their performance.
in surgery, or they were assigned to other clinical duties, and they made up this session as they were able.

Four weeks later, the final session of 10 practice trials was completed. Again, participants were given no additional instructions, but received feedback just as they did during all of the previous sticks. For the robotic group, the first and last trial were the baseline patient case and the middle 8 cases varied with increasing difficulty. After each of the 10-completed trials, participants filled out the same Practice Insertion Performance Evaluation form used in the previous practice trials.

Once the 22 practice trials were complete, all participants were required to complete two practice sessions on the entire CVC insertion procedure (which extends beyond the scope of this study but includes the needle insertion procedure) on the Manikin and their performance was evaluated, by a second-year surgical resident, using the CVC checklist that was located inside their Self-Evaluation Training Books. See Figure 5-4 for cover page of book and see Appendix E for full-details on each enclosed page. After these two full practice sessions were complete, all participants individually performed a post-test needle insertion on a manikin using the same procedures as the pre-test. Following the post-test, participants were asked to complete the same CVC Self-efficacy survey and prior experience survey as was completed at the start of the study. They were also asked to complete a Training Evaluation Survey (see Appendix B for full survey) which asked about how well they thought their training simulator helped them learn various skills.
5.2 Data Analysis and Results

To answer the research questions posed at the beginning of this chapter, statistical analyses were performed on the following items: the Central Line Self-Efficacy Survey (CLSE), the Practice Insertion Performance Evaluation Forms, and the objective score generated by the DHRT. These analyses and their results are presented in the following sections in relation to the research questions. All analyses were conducted using SPSS (v. 24.0) with an error rate of 0.05 unless otherwise specified. The remainder of this section highlights the findings with relation to the research questions.

Figure 5-4: The Self-Evaluation Training Books were filled out by participants throughout the study. For each needle insertion, participants recorded the date, a self-rating of performance, the number of attempts required to access the vessel, any mistakes made, and if feedback was provided.
**RQ1: Do surgical residents improve their CVC self-efficacy over the course of training and is this improvement dependent on the type of training provided?**

The first research question sought to understand if resident CVC self-efficacy changed from pre- to post- CVC training regardless and, if so, if this difference was due to the CVC training condition. The first hypothesis was that self-efficacy would be increased by training regardless of the resident’s training group because prior work has shown that simulator training increases resident comfort with CVC procedures\(^2\). In order to test this hypothesis, statistical analyses were performed on the pre- and post- training CLSE (Central Line Self Efficacy) Survey. As a reminder, this survey contained 14 questions which asked participants to rate their confidence in their ability to perform the skills necessary to insert a needle for the central line procedure on a 5-point Likert scale from 1- to 5 with 1 being ‘not at all confident’ and 5 being ‘very confident’ (see the 4.1.3 Metrics section in Chapter 4 for details).

The results showed that residents in the manikin training group significantly improved their self-efficacy on eleven out of the fourteen self-efficacy questions. Residents in the DHRT training group significantly improved their self-efficacy on six of the eleven self-efficacy items. When compared directly, residents in the manikin training condition were more likely to improve their self-efficacy for “placing the needle at the center of the vessel in one attempt” and residents in the robotic training condition were more likely to improve their self-efficacy for “using tactile feedback to help guide the introducer needle”.

Specifically, in order to analyze the change between responses on the pre- and post- CLSE Survey, an ordinal logistic regression was conducted with each survey question as a dependent variable and the test order (pre- or post-) as the factor, see Table 5-1 for median survey response data. For the manikin training group, the results showed that all survey questions were significantly improved from the pre- to the post-test. For the DHRT training group, all survey questions were significantly improved from the pre-to post- test with the exception of “Locating the needle on the ultrasound image” which had a median value of 4 for both the pre- and post- tests. These results support the hypothesis that self-efficacy would increase throughout the surgical residents CVC training regardless and supports prior work that simulator training has a positive effect on surgical resident comfort with CVC procedures\(^2\). This result is also supported by the
findings of the study presented in Chapter 3 which showed that the DHRT system was an effective method for increasing self-efficacy of ability to complete CVC skills.

Table 5-1: Self-Efficacy Responses

<table>
<thead>
<tr>
<th>Task Description</th>
<th>Median Pre-test</th>
<th>Median Post-test</th>
<th>p-value</th>
<th>Median Pre-test</th>
<th>Median Post-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Obtaining a clear image of the target vessel using the ultrasound machine</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.004</strong></td>
<td>4</td>
<td>5</td>
<td><strong>0.007</strong></td>
</tr>
<tr>
<td>Locating the needle on the ultrasound image</td>
<td>3</td>
<td>4</td>
<td>*0.011</td>
<td>4</td>
<td>4</td>
<td>0.192</td>
</tr>
<tr>
<td><strong>Identifying the correct site of insertion on the skin for the introducer needle</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.001</strong></td>
<td>3</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td><strong>Using tactile feedback (sense of feel) during placement of the line</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.002</strong></td>
<td>2</td>
<td>4</td>
<td><strong>0.002</strong></td>
</tr>
<tr>
<td>Using the ultrasound image/transducer, identifying the correct vessel for puncture</td>
<td>3</td>
<td>4</td>
<td>*0.021</td>
<td>4</td>
<td>5</td>
<td>0.092</td>
</tr>
<tr>
<td><strong>Using tactile feedback, identifying the correct vessel for puncture</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.002</strong></td>
<td>2</td>
<td>4</td>
<td><strong>0.003</strong></td>
</tr>
<tr>
<td><strong>Advancing the introducer needle slowly and steadily</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.001</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.006</strong></td>
</tr>
<tr>
<td><strong>Modifying the needle trajectory based on ultrasound feedback</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.004</strong></td>
<td>3</td>
<td>4</td>
<td>*0.01</td>
</tr>
<tr>
<td><strong>Identifying when the introducer needle is in the correct location</strong></td>
<td>2</td>
<td>4</td>
<td><strong>0.001</strong></td>
<td>3</td>
<td>4</td>
<td>*0.016</td>
</tr>
<tr>
<td><strong>Using tactile feedback to help guide the introducer needle</strong></td>
<td>2.5</td>
<td>4</td>
<td><strong>0.002</strong></td>
<td>2</td>
<td>4</td>
<td>*0.002</td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein in one attempt</td>
<td>2</td>
<td>4</td>
<td>-</td>
<td>2</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td><strong>Placing the introducer needle at the center of the vein in multiple attempts</strong></td>
<td>2.5</td>
<td>4</td>
<td><strong>0.002</strong></td>
<td>3</td>
<td>4</td>
<td>*0.016</td>
</tr>
<tr>
<td><strong>Conducting the entire procedure without any mistakes</strong></td>
<td>2</td>
<td>4</td>
<td><strong>&lt;.001</strong></td>
<td>2</td>
<td>4</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td><strong>Conducting the entire procedure on a CVC simulator (mannequin)</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.004</strong></td>
<td>3</td>
<td>4</td>
<td><strong>0.009</strong></td>
</tr>
</tbody>
</table>

* indicates significance at p<.01, ** indicates significance at p <.05, - indicates that a p-value could not be calculated due to a singularity
In order to determine if there were any differences between the two training groups for the pre- and post- self-efficacy surveys a Generalized Estimating Equation was conducted. The between-subject variable was participant ID, the within-subject variable was the test order (pre- training or post-training), the factors were the type of training (DHRT or Manikin) and each question on the CLSE survey. Each question was run as an ordinal logistic regression and the interaction effect between training type and response is reported in Table 5-2.

Table 5-2: Values of General Estimating Equation: Pre- and Post-SEL Survey Responses

<table>
<thead>
<tr>
<th>Factor</th>
<th>β</th>
<th>Std Error</th>
<th>95% CI</th>
<th>p-value</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obtaining a clear image of the target vessel using the ultrasound</td>
<td>0.316</td>
<td>0.9428</td>
<td>(-2.164, 1.531)</td>
<td>0.737</td>
<td>0.73</td>
</tr>
<tr>
<td>Locating the needle on the ultrasound image</td>
<td>0.672</td>
<td>1.1453</td>
<td>(-2.917, 1.573)</td>
<td>0.557</td>
<td>0.51</td>
</tr>
<tr>
<td>Identifying the correct site of insertion on the skin for the</td>
<td>1.878</td>
<td>1.5043</td>
<td>(-4.827, 1.070)</td>
<td>0.212</td>
<td>0.15</td>
</tr>
<tr>
<td>introducer needle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using tactile feedback (sense of feel) during placement of the line</td>
<td>0.604</td>
<td>2.3863</td>
<td>(-2.073, 7.281)</td>
<td>0.275</td>
<td>13.51</td>
</tr>
<tr>
<td>Using the ultrasound image/transducer, identifying the correct vessel</td>
<td>0.302</td>
<td>1.3162</td>
<td>(-2.278, 2.882)</td>
<td>0.819</td>
<td>1.35</td>
</tr>
<tr>
<td>puncture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using tactile feedback, identifying the correct vessel for puncture</td>
<td>0.514</td>
<td>2.3972</td>
<td>(-2.184, 7.213)</td>
<td>0.294</td>
<td>12.35</td>
</tr>
<tr>
<td>Advancing the introducer needle slowly and steadily</td>
<td>1.322</td>
<td>1.228</td>
<td>(-3.729, 1.084)</td>
<td>0.282</td>
<td>0.26</td>
</tr>
<tr>
<td>Modifying the needle trajectory based on ultrasound feedback</td>
<td>0.078</td>
<td>1.339</td>
<td>(-2.546, 2.703)</td>
<td>0.953</td>
<td>1.08</td>
</tr>
<tr>
<td>Identifying when the introducer needle is in the correct location</td>
<td>0.039</td>
<td>1.321</td>
<td>(-2.628, 2.550)</td>
<td>0.976</td>
<td>0.96</td>
</tr>
<tr>
<td>Using tactile feedback to help guide the introducer needle</td>
<td><strong>0.762</strong></td>
<td><strong>1.7723</strong></td>
<td><strong>(0.288, 7.235)</strong></td>
<td><strong>0.034</strong></td>
<td><strong>43.03</strong></td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein in one attempt</td>
<td><strong>1.99</strong></td>
<td><strong>1.90949</strong></td>
<td><strong>(-4.136, 0.156)</strong></td>
<td><em>0.069</em>*</td>
<td><strong>0.13</strong></td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein in multiple</td>
<td>0.649</td>
<td>1.1986</td>
<td>(-2.998, 1.700)</td>
<td>0.588</td>
<td>0.52</td>
</tr>
<tr>
<td>attempts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conducting the entire procedure without any mistakes</td>
<td>0.14</td>
<td>1.244</td>
<td>(-2.579, 2.299)</td>
<td>0.91</td>
<td>0.86</td>
</tr>
<tr>
<td>Conducting the entire procedure on a CVC simulator (mannequin)</td>
<td>0.403</td>
<td>0.9951</td>
<td>(-1.547, .164)</td>
<td>0.685</td>
<td>1.49</td>
</tr>
</tbody>
</table>

* indicates significance at p<.10, ** indicates significance at p <.05
The results showed that for the CLSE question "Using tactile feedback (sense of feel) to help guide the introducer needle" there was a statistically significant interaction effect between pre/post test and training method (p =0.034) with a beta value of 3.762 and an odds ratio of 43.03. This indicates that individuals in the robotic group were 43 times more likely than the manikin group to increase their self-efficacy rating for this measure. In addition, there was a moderately statistically significant interaction effect between pre/post test and training method (p =0.069) for "Placing the introducer needle in the center of the vein in one attempt" with a beta value of -1.990 and an odds ratio of 0.13. This indicates that individuals in the manikin group were more likely than the robot group to increase their self-efficacy rating for this measure.

These findings confirm our hypotheses that residents who were trained in the robotic group would have equal or better increases in self-efficacy. They also confirm prior findings that providing feedback after good or accurate performance increases self-efficacy82,87 and providing learners with augmented feedback (or feedback not normally present) can help engage cognition during the learning process93. However, it is important to note that while the DHRT was designed to provide detailed specific, and objective feedback which has been shown to improve skill gains and retention75 this did not show any advantage over feedback provided from the manikin. In fact, self-confidence for placing the needle at the center of the vein was lower. The impacts of this are discussed further the Discussion section of this chapter.

RQ2: How do perceived ratings of performance change throughout training and is this dependent on the training method?

The previous research question focused on identifying how the CVC training method affected residents’ CVC self-efficacy. The current research question focuses instead on understanding differences in self-perception of performance during training. Specifically, this question sought to address how participant self-ratings of performance compared across training groups over the duration of training. It was hypothesized that perceptions of performance, as measured through self-ratings, would increase throughout training because research has indicated that motor skills and self-perception of those skills improve with training82,87. It was hypothesized that the robotic training group would have lower self-ratings of performance compared to the manikin group because the
robotic group was provided with consistent objective feedback on performance. This is because prior research has indicated that novices have a low ability to judge skill performance and are over-confident\textsuperscript{55,56}, but when provided with realistic feedback they may be better equipped to accurately assess their own performance\textsuperscript{96}. The results showed that self-rating scores increased over time for both the manikin training condition and the robotic training condition. In addition, participants in the manikin training condition had higher self-rating scores across all trials, but recorded less feedback than participants in the robotic training condition.

Specifically, in order to answer this research question, two separate analyses were conducted to investigate: (1) how self-rating of performance changes with each training group, and (2) the differences between training groups for self-ratings. These analyses utilize the self-ratings from the Self-Evaluation Training Books where participants were asked to rate their performance on a scale of 1-5 with 1 being “very poor” and 5 being “very good” as shown in the example page of the Self-Evaluation Training Book in Figure 5-1.

In order to understand how self-ratings changed within each training group, a Wilcoxon Signed Ranks test was run to compare self-ratings on the first and last needle insertions for both groups with 'self-rating' as the dependent variable and ‘insertion number’ as the independent variable. Assumptions were checked and the distributions were approximately symmetrically shaped. The results revealed a statistically significant (p < .01) increase in self rating for both training groups as indicated in Table 5-3. This shows that for both groups, participants felt they had improved their CVC insertion performance.

<table>
<thead>
<tr>
<th></th>
<th>Insertion 1</th>
<th>Insertion 22</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robotic</td>
<td>Median 3</td>
<td>Median 4</td>
<td>2.919</td>
<td>*0.004</td>
</tr>
<tr>
<td>Manikin</td>
<td>Median 3</td>
<td>Median 5</td>
<td>2.762</td>
<td>*0.006</td>
</tr>
</tbody>
</table>

* indicates significance at p<.01,

Once it was identified that there were increases in self-ratings for both conditions a Generalized Estimating Equation was run to determine if there were any differences between the two training groups for increases in their self-ratings of performance.
Specifically, the between-subject variable was participant ID, the within-subject variable was the trial number, the factors were type of training (robotic or manikin) and self-rating of performance on scale of 1 (very poor) to 5 (very good). The ordinal logistic regression showed a statistically significant main effect for training method ($p < 0.000, \chi^2 = 20.517$) and trial number ($p < 0.000, \chi^2 = 1020.276$). These results indicate that in general the manikin group rated their performance higher than the robotic group did throughout the trials, see Figure 5-2.

In addition, the results revealed that ratings for trials 1, 2, 3, 4, 13, and 16 were statistically significant, indicating that the rating for those trials were each statistically different from mean rating of the rest of the trials, see Table 5-4. Importantly, trials 1 and 2 were the first attempts at the procedure with feedback, so it was not surprising that self-ratings were lower on these days because the findings from the study conducted in Chapter 4 indicate that a learning curve was present for novices when using the DHRT system. On the other hand, trials 3 and 4 were the first two trials of a second training session and trial 13 was the first trial of the third training session. Thus, it is not surprising that due to degradation in learning between practice sessions that the mean ratings of performance would be significantly lower. These findings are supported by prior research, which indicates that skill retention declines over time for CVC procedures\textsuperscript{122}. In addition, trial 16 was statistically significantly lower as well. Figure 5-5 shows that the mean rating for performance on the manikin was similar to the surrounding trials, but that self-ratings for the robotic group were visibly lower. This suggests that possibly the patient case had an effect on performance and self-rating. This is further discussed in the Discussion section of this chapter.
In addition, there was a statistically significant interaction effect for the entire model between trial and training group (p < 0.000, $\chi^2 = 196.273$); which means that for each individual trial there was no significant difference between how the manikin and robotic group rated their performance on each individual trial (e.g. Trial 1). However, when all 22 trials were considered together, the way that the ratings changed over time was different for each training group. This may be impacted by the type of feedback received by participants which varied in each group varied by the content, method of delivery, and frequency – all of which have been shown to influence learning.

In order to understand if these differences in self-ratings of performance were potentially impacted by the amount and type of feedback participants were receiving, content analysis was performed on the responses to “List any errors you made:_____” recorded in the Self-Evaluation Training Books after each insertion trial, see Figure 5-1. In order to do this, the responses were transcribed and analyzed by two independent raters using combined inductive and deductive content analysis in NVivo v11.4.0. Specifically, each feedback item from the Learning Feedback Screen, with the exception of overall score, was made into a node in NVivo. This resulted in the following 6 nodes:
angle, aspiration, distance to center of the vein, number of insertions, and arterial puncture. Next, two researchers verbally discussed each node until both felt satisfied that they had a mutual understanding of each item. Each rater then, independently, rated 7 of the participant books and added any nodes they felt were frequently mentioned. Together the raters then reevaluated the existing nodes to form a consensus for which nodes were a good representation of the feedback being provided. The following 3 nodes were added: ‘external landmark’, ‘left versus right hand’, and ‘ultrasound monitor’. Once the coding scheme was set an interrater reliability (weighted Kappa) of 0.72 was achieved, the remaining 19 Self-Evaluation Participant Books were coded, see Table 5-4 for representative coding schema.

Of the 372 errors recorded in the Self-Evaluation Participant Books, 144 were reported from the manikin group and 228 from the robotic group, see table 5-5 for breakdown in errors recorded. In order to determine whether an equal number of participants from each of the training conditions reported each type of errors during their CVC training, Chi-squared goodness of fit tests were conducted. The results indicated that the number of errors were not equally represented by participants in the study, see table 5-5 for full lists of chi-squared results. Specifically, the robotic group reported a higher number of total errors ($\chi^2 (1) = 18.97, p < 0.001$), errors in the angle of insertion ($\chi^2 (1) = 14.52, p < 0.001$), errors in the needle tip’s final distance to the center of the vessel ($\chi^2 (1) = 19.15, p = 0.002$), and errors in having multiple insertion attempts ($\chi^2 (1) = 23.06, p = 0.005$). There was also a marginally significant finding for errors in aspirating during the insertion attempt ($\chi^2 (1) = 3.45, p < 0.063$). On the other hand, participants in the manikin group reported a higher number of errors in pressure or torque during the insertion attempts ($\chi^2 (1) = 8.00, p = 0.005$).

These findings confirm the hypothesis that perceptions of performance, as measured through self-ratings, would increase throughout training. This is supported by prior research that indicates self-perception of skill performance improves with training$^{82,87}$. The findings also support the hypothesis that the robotic training group would have lower self-ratings of performance compared to the manikin group. This is possibly because the robotic group was provided with consistent objective feedback on performance. The content analysis of the feedback recorded by participants indicates that participants in the robotic training group received more feedback than those in the manikin training group. These findings support prior research which has indicated that novices have a low ability to judge skill performance and are over-confident$^{55,56}$, but
when provided with realistic feedback they may be better equipped to accurately assess their own performance\(^9\).

Table 5-4: Transcribed Errors from the Self-Evaluation Training Book by Participants AMON08 (Manikin) and CYTY07 (Robotic) and the Code of the Item From the Content Analysis.

<table>
<thead>
<tr>
<th>Insertion Number</th>
<th>Manikin Group Participant</th>
<th>Robotic Group Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>NVIVO Code</td>
<td>Response</td>
</tr>
<tr>
<td>1 probe orientation, aspire as you insert needle</td>
<td>-Probe, -Aspiration</td>
<td>2 insertions</td>
</tr>
<tr>
<td>2 NA</td>
<td>-</td>
<td>failed to access, far from center</td>
</tr>
<tr>
<td>3 went past vein</td>
<td>-Distance to center of the vein</td>
<td>2 insertions</td>
</tr>
<tr>
<td>4 NA</td>
<td>-</td>
<td>2 insertions, pulled out of vein at end</td>
</tr>
<tr>
<td>5 NA</td>
<td>-</td>
<td>pulled out after removing probe and had to re-access</td>
</tr>
<tr>
<td>6 NA</td>
<td>-</td>
<td>still pulling out when removing probe, 2 insertions</td>
</tr>
<tr>
<td>7 NA</td>
<td>-</td>
<td>pulled out when removing probe removed</td>
</tr>
<tr>
<td>8 NA</td>
<td>-</td>
<td>didn't pull out when removing probe but required 2 attempts to access</td>
</tr>
<tr>
<td>9 NA</td>
<td>-</td>
<td>2 attempts to access</td>
</tr>
<tr>
<td>10 NA</td>
<td>-</td>
<td>2 attempts to access</td>
</tr>
<tr>
<td>11 NA</td>
<td>-</td>
<td>attempts because not close to center and losing access</td>
</tr>
<tr>
<td>13 didn’t aspirate initially, went through vein, but was able to troubleshoot</td>
<td>-Aspiration -Distance to the center of the vein</td>
<td>angle too steep</td>
</tr>
<tr>
<td>14 NA</td>
<td>-</td>
<td>pulled needle out</td>
</tr>
<tr>
<td>15 NA</td>
<td>-</td>
<td>shallow angle</td>
</tr>
<tr>
<td>16 NA</td>
<td>-</td>
<td>multiple attempts</td>
</tr>
<tr>
<td>17 NA</td>
<td>-</td>
<td>2 attempts</td>
</tr>
<tr>
<td>18 NA</td>
<td>-</td>
<td>hit carotid, multiple passes</td>
</tr>
<tr>
<td>19 NA</td>
<td>-</td>
<td>2 attempts</td>
</tr>
<tr>
<td>20 NA</td>
<td>-</td>
<td>3 attempts? fairly certain hit vein and stayed in</td>
</tr>
<tr>
<td>21 NA</td>
<td>-</td>
<td>2 attempts</td>
</tr>
<tr>
<td>22 NA</td>
<td>-</td>
<td>2 attempts, close to center</td>
</tr>
</tbody>
</table>
Table 5-5: Number of Errors Recorded by Each Training Group in the Self-Evaluation Training Book and the results of the Chi-Squared test.

<table>
<thead>
<tr>
<th>Error Recorded</th>
<th>Manikin</th>
<th>Robot</th>
<th>$\chi^2$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>21</td>
<td>54</td>
<td>14.52</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Aspiration</td>
<td>18</td>
<td>31</td>
<td>3.45</td>
<td>0.063</td>
</tr>
<tr>
<td>Distance to Center of Vein</td>
<td>43</td>
<td>76</td>
<td>9.15</td>
<td>0.002</td>
</tr>
<tr>
<td>External Landmarks</td>
<td>5</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Left vs. Right Hand</td>
<td>6</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mismatch with Probe</td>
<td>18</td>
<td>16</td>
<td>0.12</td>
<td>0.732</td>
</tr>
<tr>
<td>Number of Insertions</td>
<td>3</td>
<td>31</td>
<td>23.06</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Pressure or Torque</td>
<td>15</td>
<td>3</td>
<td>8.00</td>
<td>0.005</td>
</tr>
<tr>
<td>Arterial Puncture</td>
<td>0</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>15</td>
<td>13</td>
<td>0.143</td>
<td>0.705</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>228</td>
<td>18.97</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Bold indicates significance at p < 0.01, italics at p < 0.10.

**RQ3: How do perceived ratings of performance on the robotic trainer relate to objectively calculated performance scores in the DHRT system?**

The third research question sought to address how participant self-ratings of performance during training compared with the objective score generated by the DHRT system. The self-rating of performance in this case was the rating that was put into the DHRT system prior to receiving feedback (see Figure 5-6). It was hypothesized that there would be a weak positive correlation between self-ratings of trial performance and scores generated by the DHRT system because providing individuals with feedback during training has been shown to *improve* their ability to accurately assess their own performance\(^96\). However, previous research has also indicated that even experts may have
a low ability or no ability to judge their own skill performance\textsuperscript{55,56,97}. The results showed that there was a positive, weak, but post statistically significant relationship between perceived self-ratings of performance and objectively calculated scores in the DHRT system.

Specifically, in order to answer this research question, a Linear Regression analysis was conducted in order to determine if the self-ratings of performance obtained before getting feedback from the DHRT system were able to predict the objective score generated by the DHRT system. Specifically, for this model, the dependent variable was the objective score generated by the DHRT system and the independent variable was the participant self-rating of performance indicated in the DHRT system before feedback was received. The regression analysis was computed only after the assumption of a monotonic relationship was confirmed through a visual evaluation of a scatter plot.

![Assessment of a Monotonic Relationship](image)

Figure 5-6: Scatterplot used to assess for the existence of a monotonic relationship between self-rating of performance and the score generated by the DHRT System. The line indicates that both self-ratings and the score generated by the DHRT increase at a constant rate.

The results show that self-ratings of performance on the DHRT system could significantly predict the objective score of the DHRT system, \( R^2 = 0.223, F (1,285) = 81.294, p < 0.001 \); adjusted \( R^2 = 0.220, B = 0.109 \). These results indicate that participants
who used the DHRT system were able to predict their objective performance score on the DHRT system. This supports our hypothesis that there would be a positive correlation between self-ratings of performance and objective measures of performance and it also supports prior research that indicates novices can judge skill their skill performance when provided with feedback that directs their attention to performance cues96.

**RQ4: How do perceptions of the effectiveness of the training simulator differ between simulator fidelity?**

The final research question sought to identify if there were any differences between training groups when participants were asked to evaluate how effective the training they received was for increasing the skills required to complete a CVC procedure. It was hypothesized that participants in the robotic training group would rate their training method as equally effective or more effective than participants in the manikin training group because prior research has shown that VR training systems can be effective for teaching needle insertion procedures31. Additionally, the results presented in Chapter 3 demonstrate that the individuals trained on the DHRT system found it to be as effective as being trained on the manikin and the results from Chapter 3 also revealed a marginally significant difference between the conditions for the question “helped me understand how to modify CVC insertion procedures based on patient anatomy” with participants in the robotic training group rating this item higher than those in the manikin training group. The results showed that perceptions of the effectiveness of the training simulator did not differ between simulator fidelity.

Specifically, in order to test this hypothesis, an analysis was conducted on the Training Evaluation Survey (TES) in order to compare residents’ feelings about the utility of the training methods for learning CVC procedures. See Appendix B for the full survey. Specifically, a Mann-Whitney U-test was computed with the dependent variables being the response to each item on the TES and the independent variable being the method of training (manikin or robot only). The results failed to identify any significant differences between responses on each item of the TES in the two conditions, see results in Table 5-6. These results support the hypothesis that participants in the robotic training group would rate the effectiveness of the DHRT system equal to or higher than the ratings provided by the manikin training group for the manikin simulator. This supports
prior research which has shown that VR simulators can provide sufficient realism for teaching ultrasound guided needle insertion procedures\textsuperscript{31}. However, these results did not replicate the findings from Chapter 3 which indicated that DHRT simulator helped participants understand how to modify CVC insertion procedures based on patient anatomy. This difference may be because the manikin participants in this study were not given any patient profiles and therefore may not have considered the effects of patient anatomy while practicing the procedure.
The goal of this chapter was to determine how self-efficacy and self-ratings of performance were affected by training on either the DHRT or manikin system. This is

<table>
<thead>
<tr>
<th>The method of training I received:</th>
<th>Median Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manikin</td>
</tr>
<tr>
<td>…made me sensitive to how patient anatomy impacts the force required to advance needles in the human body</td>
<td>4</td>
</tr>
<tr>
<td>…made me sensitive to patient anatomies' impact on vessel location in the body</td>
<td>4</td>
</tr>
<tr>
<td>…helped me understand how to modify CVC insertion procedures based on patient anatomy</td>
<td>4</td>
</tr>
<tr>
<td>…helped me learn where to insert the introducer needle during the procedure</td>
<td>4</td>
</tr>
<tr>
<td>…helped me learn to distinguish the vein from artery on the ultrasound image</td>
<td>4</td>
</tr>
<tr>
<td>…helped me learn how and when to modify my needle's trajectory during the insertion</td>
<td>4</td>
</tr>
<tr>
<td>…helped me correctly identify when I had successfully inserted the needle</td>
<td>4</td>
</tr>
<tr>
<td>…helped me understand how to use tactile feedback (sense of feel) during placement of the line</td>
<td>4</td>
</tr>
<tr>
<td>…was an effective method for learning the CVC process</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 5-6:** Results of the Mann-Whitney U-test which failed to indicate any significant differences in the training effectiveness between the Manikin and Robotic training conditions.

5.3 **Discussion**

The goal of this chapter was to determine how self-efficacy and self-ratings of performance were affected by training on either the DHRT or manikin system. This is
important because although there are links between self-efficacy and skill performance\textsuperscript{48}, this relationship may diminish when performing the task in novel situations\textsuperscript{54,55}. The main findings of the study were as follows: (1) Participants in both the manikin and robotic condition showed statistically significant improvements in their CVC insertion skill self-efficacy over the course of their training but participants trained on the DHRT system were more likely to have increased confidence in their ability to use tactile feedback to guide the introducer needle over the course of the training while participants in the manikin condition were more likely to increase their confidence in their ability to place the needle in the center of the vessel; (2) Participants in both training groups increased their self-ratings over the course of their training but participants in the manikin group rated their self-performance higher than participants in the DHRT system over the course of the training; (3) Self-ratings of performance were significantly lower for trials 1,2,3,4,13 and 16; (4) Participants who used the DHRT system reported significantly more errors during their training; (5) Participants who used the DHRT system were able to predict their objective performance score on the DHRT system; and (6) Participants in both groups viewed their training simulator as an effective method for improving CVC insertion skills.

These results are encouraging because both training groups showed improvements in resident self-efficacy which has been shown to relate to skill performance\textsuperscript{54,88}. While the analysis of the Central Line Self-Efficacy survey revealed that training on either the manikin or the DHRT system improved self-efficacy for CVC insertion skills, participants trained on the DHRT system showed larger improvements in their confidence in their ability to use tactile feedback to guide the introducer needle. On the other hand, participants trained on the manikin had larger gains in confidence in their ability to place the needle in the center of the vessel. This finding is interesting because the DHRT system provides specific feedback on the closeness of the needle to the center of the vein after each trial, and this information is hard to provide objective feedback on for the manikin training system. As such, residents are unlikely to receive feedback on this skill during training and may have a false sense of confidence in this ability. However, future work is needed to explore why this is occurring and how accurate these self-efficacy gains are with actual performance on the manikin trainer. Regardless, the results presented here indicate that the DHRT system is at least as effective as the manikin for increasing resident self-confidence with CVC insertion skills confirming
prior research that found that self-efficacy can be improved through the practice of skills\textsuperscript{20}.

In addition, the results showed that participants in both the manikin and robotic training groups increased their self-ratings over the course of their training. Specifically, individuals in the manikin group, on average, rated themselves higher than the participants in the robotic group. Despite this, the participants in the manikin group recorded approximately half the amount of errors as those in the robotic group. This is important because prior research has suggested that providing individuals with specific feedback about their performance, and specific feedback about how to evaluate their performance, can increase the accuracy of their self-evaluations\textsuperscript{96}. This suggests that the participants in the robotic group may have been better at evaluating their performance than the individuals in the manikin group, which may indicate that participants in the manikin group were overconfident in their abilities. In addition, this finding supports prior literature that has indicated that surgical novices are overconfident in their performance\textsuperscript{98}.

This finding is also supported by the content analysis of the Self-Evaluation Training Book, which revealed that participants in the robotic group recorded nearly double the amount of errors based on feedback on their performance compared to the manikin group. These errors specifically included distance to the center of the vein, angle of insertion, number of attempts, and continuous aspiration, all of which relate to an awareness of where the tip of the needle is in relation to the vessels. This relates to prior research which has indicated that, when provided with specific feedback on cues that are indicative of performance, self-evaluation skills will improve in accuracy\textsuperscript{96}. This suggests that participants in the robotic training group may be more adept at evaluating their performance when compared to the manikin group. In addition surgical expertise is developed through deliberate practice and feedback; and the opportunity to incorporate the feedback is critical for developing expertise in surgery\textsuperscript{75}. This is likely because evaluating performance and noticing errors may result in greater attention allocation to the task and therefore lead to increased learning and skill retention\textsuperscript{94,95}. This is important because if residents can improve their performance on these skills, they may be more adequate in determining where their needle tip is located and therefore less likely to cause a pneumothorax (puncturing a lung), or a hematoma (excessive bleeding) caused by arterial puncture, both of which are common complications for IJ central line placement\textsuperscript{37,48}. 
This is especially interesting when considering the fact that the robotic training group had consistently lower self-ratings of performance even though the analysis of the CLSE revealed that both groups were equally confident in their ability to perform CVC skills. In addition, both groups felt that their training system was an effective method of teaching the necessary skills for CVC placement. These findings suggest that there may be a discrepancy between how well novices think they are performing and how well they are actually performing. This is supported by the finding that participants’ self-ratings of performance were not strongly correlated to the score generated by the DHRT. However, this is not surprising because prior research has indicated that novices have a low ability to judge skill performance\textsuperscript{55,56}. Specifically, research has found that there is little or no relationship between self-assessment and objective measures of performance\textsuperscript{97}. The ability to accurately assess performance can be improved by providing feedback on cues that are indicative of performance\textsuperscript{96} and the DHRT system was designed to give feedback on objective measures that were indicative of performance. These findings, combined with prior research\textsuperscript{98} strongly suggest that individuals trained in the manikin group were over confident in their skills. While these results indicate that that self-efficacy and self-ratings are not good measures of performance\textsuperscript{67} because research has shown that self-efficacy can impact skill gains\textsuperscript{89} and performance\textsuperscript{57}.

Finally, the results also showed that regardless of the training condition, self-ratings were significantly lower for trials 1, 2, 3, 4, 13, and 16 over the course of the training. Importantly, trials 1-4 and 13 were the first practice insertions for training sessions 1, 2 and 3, respectively. Thus, this finding is not surprising since this was the first practice needle attempts, and the break and practice can affect perceptions of performance. This finding supports prior research that has shown that skill performance declines over time if not regularly practiced\textsuperscript{21}. On the other hand, trial 16 was not one of the first insertion attempts or a trial that occurred after a break. It is important to note that while not statistically significant, this self-rating of performance was lower for the robotic group compared to the manikin group. The scenario for this trial was scenario 11 which represented a patient with very thin skin and shallow vessels that were nearly vertical, see Appendix A. Of the 13 robotic participants, 1 participant punctured the artery and 3 participants punctured through the back wall of the vein during this trial meaning that this trial may have been more difficult than some of the other cases presented to the robotic group, see Pepley et. al\textsuperscript{123} for detailed findings. Thus, the
findings for trial 16 may be related to the case presented. However, future work is needed to further investigate the impact of the patient case on perceived performance.

While these results are promising for the use of the DHRT system, there are several limitations to this study. Most notably, the pre- and post- tests were conducted on the same manikin that was used for the manikin training group. And prior research suggests that self-efficacy may decrease when transferring skills to a novel situation. This is encouraging for the use of the DHRT as a training method since participants trained using this device were equally confident in their ability to perform the procedure but it does limit the conclusions that can be drawn about participants in the manikin group. Additionally, because the findings indicate that novices may not be able to accurately judge their performance the analyses reported in this chapter should be supported by an analysis of actual skill performance. In order to account for this, Chapter 6 focuses on understanding the relationship between skill gains and training methods.
Chapter 6
Measuring Surgical Residency Skill Gains in Central Venous Catheterization Training: An Exploration in Simulator Fidelity

Publication: Mary Yovanoff, David Pepley, Katelin Mirkin, Jason Moore, David Han and Scarlett Miller. “Measuring Surgical Resident Skill Gains in Central Venous Catheterization Training: An Exploration in Simulator Fidelity” To be Submitted to the Journal of Academic Medicine, July 2017.

The previous chapter highlighted the impact of the type training type (manikin or DHRT) had on self-efficacy and self-rating of performance. However, it did not consider how these training methods impacted observer-based and objective-based measures of CVC skill performance or how these performance measures changed over the course of training. This is important because there is no single standard for evaluating skill competency in CVC placement. However, most studies evaluating CVC skills training use checklists filled out through supervisor-evaluations such as The Global Rating Scale (GRS) Assessment tool (a 9-item Likert scale assessment intended for use generally across multiple skills) or a rigorously developed 29 item binary checklist specific to CVC skills. The vast majority of these evaluation methods are used to compare trainee skills to some baseline and evaluate if they are good enough to ‘pass’ and perform the procedure in clinical situations.

While checklists are useful for standardizing practices, they may not be appropriate to use during skills training because they can encourage “teaching to the test”. In addition, checklists can result in observers simply checking off boxes rather than providing helpful feedback which is a critical part of the learning process. This suggests the need to distinguish between evaluation tools and training tools. There has been some effort to develop tools more appropriate for training rather than just skill evaluation, such as the Imperial College Surgical Assessment device (ICSAD), through the use of motion analysis. These assessments use objective measures to evaluate skill performance through metrics such as path length, time to complete, steadiness of motion, velocity, total amount of movement, and overall efficiency of motion. Importantly, the use of electromagnetic sensors to identify motion paths during CVC training show promise as a valid and objective assessment of skills and these metrics can be captured using VR simulators.
The development of objective measures of performance can also enhance the quality of feedback that can be provided to novices during training. Research has shown that providing novices with detailed and appropriate feedback is a critical part of the learning process and the opportunity to incorporate that feedback is crucial to the development of expertise in surgery. The DHRT system was designed to objectively measure performance on critical metrics and provide feedback on these metrics. It has been suggested that the evaluation of competency of a trainee at a surgical procedure should include a combination of the trainee’s self-perception of ability, an evaluator’s perception of their ability, and an objective skills based evaluation of their ability. Therefore, the study presented in this chapter was designed to explore the effectiveness of the DHRT system and learning interface for increasing CVC skills using both objective and subjective measures as compared to standard manikin training methods, and then identify if a relationship exists between these objective and subjective measures of performance.

6.1 Methodology

The goal of this chapter was to identify how subjective and objective measures of CVC skills performance changed across the surgical residency training program, how this change was impacted by the type of training provided and how these subjective and objective measures complemented or contradicted one another. This is important because most evaluation methods for CVC training utilize some form of a skills checklist which may not provide sufficient feedback on technical skills for developing expertise in CVC insertion procedures. Specifically, the study was developed to answer the following research questions (RQ):

RQ1. How do subjective observations of performance change throughout training and is this dependent on the training method? Specifically, this question sought to address how a subjective resident evaluation of surgical resident’s performance using the IJ CVC skills checklist changed from pre- to post-training and if these changes were different between the training groups. It was hypothesized that there would be no differences in observed ratings on the IJ CVC skills checklist because prior research has indicated that checklists may not be granular enough to detect changes in performance.
*RQ2.* How do resident CVC skills quantified through motion analysis change through CVC training and is this dependent on the training method employed? Specifically, this research question sought to address how the total path length of the needle tip, time to complete the CVC insertion, average angle of insertion, velocity, and jerk changed before and after CVC training and if there was a difference between training groups (manikin and robotic). It was hypothesized that these objective measures of skill gains would increase with training and the skill gains for the robotic training group would be greater than or equal to the manikin training group. This was hypothesized because the study reported in Chapter 3 found that the DHRT was able to increase performance on surgical skills as have studies on the use of virtual reality (VR) simulators in other areas of surgical skills training\(^{58}\).

*RQ3.* Is there a relationship between objective and subjective performance improvements? Specifically, this research sought to identify if there was any correlation between performance on the pre- or post- test as measured by the IJ CVC skills checklist and performance on any of the 5-objective metrics measured through motion tracking: total path length of the needle tip, time to complete the CVC insertion, average angle of insertion, velocity, and jerk. It was hypothesized that there would not be a correlation between performance on the IJ CVC skills checklist and any of the objective measures of performance because prior research has indicated that skills checklists may not effectively evaluate skills needed for CVC procedures\(^{74}\) but that motion tracking\(^{112}\) can be used to identify improvements in surgical skill performance.

### 6.1.1 Participants and Procedures

The analyses presented in this chapter were conducted with same participants and procedures described in Chapter 5. The procedures for this study can thus be found in the Procedures section of Chapter 5. As indicated in Chapter 5, three participants (all male) were unable to finish the study due to constraints on their residency training. Of the three participants, one was in the robotic training group and two were in the manikin training group. Their data was included in the analysis when available. Data was not available for the post-test IJ CVC skills test or Post-test motion tracking. Additionally,
one participant from the manikin training group only completed 12 out of the 22 required needle sticks.

It is important to note for the current chapter the details of the needle tracking system. Specifically, the Ascension 3D Guidance trakSTAR (Sheldburne, VT) electromagnetic position tracking system works by using a transmitter to generate a magnetic field. This field is then measured by an electromagnetic sensor. The orientation of this sensor effects the field measurements and an output of the x-position, y-position, z-position, and probe angle can be calculated. In order to track the position of the syringe, the electromagnetic sensor was attached to the syringe at the 0 mL mark and oriented with the tip pointing towards the needle. The needle tip was located 9cm in the axial direction from the end of the sensor, allowing its location to be calculated using the location and orientation of the sensor. A sample rate of 80 Hz was used for these tests. The set-up is shown in Figure 6-1.

![Image](image.jpg)

**Figure 6-1**: Attachment of the motion tracker (indicated by the red circle) to the syringe for the pre- and post- tests which was attached to the Ascension 3D Guidance trakSTAR electromagnetic position tracking system.
6.1.2 Metrics

The five objective motion tracking metrics, developed in Chapter 4, were computed from the motion tracking of the needle pathway including: average angle of insertion, path length, velocity, jerk, and total time to complete. These metrics were included because path length, time to complete and velocity were identified in Chapter 4 as metrics that distinguished expert performance on the DHRT trainer. However, it is important to note that although distance to the center of the vein was also identified as a criterion that distinguished expert performance in Chapter 4, it was not used in the current chapter since there is no reliable method of calculating this metric on the manikin. On the other hand, Average angle of insertion was analyzed in the current chapter because of how frequently verbal feedback was shown to be provided on the manikin trainer, (see Chapter 2), because of the specific feedback provided in the DHRT personalized learning system, as discussed in Chapter 2, and because of its presence on the IJ CVC checklist. Finally, jerk, the derivative of acceleration with respect to time, was used in place of standard deviation of the deviations. This metric was chosen because it has been utilized as method for calculating smoothness of motion in surgical procedures and therefore the results can be more broadly compared with prior findings. Velocity and jerk were calculated for each vector using the Savitzky-Golay method which resulted in an x-, y-, and z-velocity and an x-, y-, and z-jerk.

It is also important to note that for each participant trial each of the data sets was trimmed to include only the motion path while the needle was being inserted into the ‘skin’ of the manikin. To do this, two plots were created from the raw data including a chart with the Pythagorean distance versus time and a chart with the needle angle versus time, see Figure 6-2. The combination of these two plots showed a distinct point in time when the needle slowed down and rested at a precise angle before entering the manikin. Two researchers independently cut the data at the overlapping point in time when the data leveled out. Once the data were trimmed, time to complete, average angle, path length, average velocity, and average jerk were calculated. The following paragraphs outline how and why each of these metrics were calculated.
**Needle Insertion Start Point**

![Graph showing Pythagorean Distance and Needle Angle over time](image)

Figure 6-2: Plot used to identify the point of needle insertion. The Pythagorean Distance is shown in Blue and the Angle of the Needle tip is shown in Orange. In this case, data was cut at Time = 9.708 sec

*Time to complete* was calculated by subtracting the start time from the end time. This was calculated as a metric because research has found that experts are faster than novices in completing basic surgical skills.\(^{103-105,107}\)

*Average angle* was calculated by taking the average of the angles between each position point. This was used as a metric because of how frequently verbal feedback was shown to be provided on the manikin trainer, (see Chapter 2), because of the specific feedback provided in the DHRT personalized learning system, as discussed in Chapter 2, and because of its presence on the IJ CVC checklist.

Needle tip *path length* was chosen as a metric because prior research has indicated that overall path length of hand motions can distinguish between expert and novice performance.\(^{102,105}\) However, path length of the needle tip in CVC insertion procedures has not been studied to date. Thus, path length was calculated as a summation of movement in the X, Y and Z coordinate system based on the points recorded by the DHRT system using the Euclidian distance formula.
Average velocity of the needle tip was chosen as a metric because prior research has indicated that the velocity of hand movements can distinguish between expert and novice performance \(^{107,108}\). However, velocity of the needle tip has not been studied to date in CVC procedures. Average velocity was defined as the average velocity of the needle tip when it was below the skin surface. This was calculated based on time required to travel the distance between each set of points recorded by the DHRT system and divided by that distance.

6.2 Data Analysis and Results

In order to answer the research questions posed as part of this chapter, statistical analyses were performed on the IJ CVC checklist and each of the five motion path metrics: total path length, average velocity, average jerk, and time to complete. These analyses and their results are presented in the following sections in relation to the research questions. All analyses were conducted using SPSS (v. 22.0) with an error rate of 0.05.

RQ1: How do subjective ratings of performance change throughout training and is this dependent on the type of training received?

The first research question sought to address how a subjective resident’s evaluation of a surgical resident’s performance using the IJ CVC skills checklist changed from pre- to post- training and identify if these changes were different between training groups. Specifically, it was hypothesized that there would be no differences in observed ratings on the IJ CVC skills checklist from pre-test to post-test or between groups because the findings from Chapter 3 indicated that the IJ CVC checklists may not be granular enough to detect changes in performance. In order to answer this research question, a Generalized Estimating Equation was run to determine if there were any differences between the two training groups for the pre- and post- IJ CVC skills checklist. The between-subject variable was participant ID, the within-subject variable was the test order (pre-test or 1\(^{st}\) attempt of post-test), and the factors were type of training (DHRT or Manikin) and each question on the checklist. Each question was run as a separate binary
logistic regression and the interaction effect between training type. However, only two items were compared between the pre-and post-tests since Selecting the appropriate site for venipuncture, Correctly distinguishing artery and vein, and Confirming entry by aspiration were always marked ‘yes’ on both the pre- and post- test checklist. The remaining analysis failed to reveal a statistically significant difference between groups or from pre- to post- test for Selecting the correct ultrasound probe and using it in the correct orientation (p = 0.695), and Inserting the needle at a 30-45 degree angle (p = 0.824).

The item Number of Insertion attempts from the IJ CVC checklist was computed separately from the above analysis as it was not binary. However, it did not meet the assumptions necessary to be computed in an ANOVA so it was also run as a Poisson Loglinear regression (intended for use with counts). The results were moderately statistically significant (p = .080) from pre- to post- test. The beta value was .480 and the odds ratio was 1.6, indicating that scores on the pre- test were 1.6 times more likely than the post- test to be 1 count (insertion) higher. However, there was not a significant interaction effect between training group and test (p = .615).

These results support the notion that the IJ CVC checklist may not be granular enough to detect changes in performance and support the findings from Chapter 3. For example, multiple participants inserted the needle towards the head of the manikin, as shown in Figure 6-3, but still received an 80% on the CVC Skills Checklist as shown in Table 6-1.
Figure 6-3: Participant performing a pre-test on the manikin and inserting the needle towards the head, which is the incorrect direction.
Table 6-1: Pre-Test Results for Participant IELE05

<table>
<thead>
<tr>
<th>CVC Checklist Item</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to complete procedure (min)</td>
<td>3</td>
</tr>
<tr>
<td>Selects appropriate site for venipuncture (y_1,2_2)</td>
<td>1</td>
</tr>
<tr>
<td>Selects correct ultrasound probe and applies in appropriate orientation</td>
<td>1</td>
</tr>
<tr>
<td>Obtaining a clear image of the target vessels using the ultrasound machine</td>
<td>1</td>
</tr>
<tr>
<td>Correctly distinguishes artery and vein: demonstrates compressibility of vein</td>
<td>1</td>
</tr>
<tr>
<td>Inserts introducer needle at 35-45 degree angle from the skin</td>
<td>1</td>
</tr>
<tr>
<td>Locating the needles position on the ultrasound image</td>
<td>1</td>
</tr>
<tr>
<td>Advancing the introducer needle slowly and slowly</td>
<td>1</td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein</td>
<td>1</td>
</tr>
<tr>
<td>Number of attempts</td>
<td>1</td>
</tr>
<tr>
<td>Confirms vessel entry by aspiration of blood</td>
<td>0</td>
</tr>
<tr>
<td>Conducting the entire procedure without any mistakes</td>
<td>0</td>
</tr>
<tr>
<td>Average Score (not including number of insertions or time to complete)</td>
<td>0.8</td>
</tr>
<tr>
<td>Please provide rationale or notes about your ratings or complications in the procedure</td>
<td>inserted needle towards head; did not aspirate; wanted to look for flow on US</td>
</tr>
</tbody>
</table>

RQ2: How do resident CVC skills quantified through motion analysis change through CVC training and is this dependent on the training method employed?

The second research question sought to address how skill gains, as measured by the quantitative measures of the motion of the needle during the pre- and post- test
assessments changed throughout the CVC skills residency training and if there were any difference between the training conditions (manikin and robotic) and if there was a difference between training conditions. It was hypothesized that these objective measures of skill gains would increase with training and the skill gains for the robotic training group would be greater than or equal to the manikin training group. This was hypothesized because the study reported in Chapter 3 found that the DHRT was able to increase performance on surgical skills. This hypothesis is also supported by research that has shown that the use of virtual reality (VR) simulators can increase surgical skills training\(^5\). The results showed that participants in both training groups reduced their path length and total time to complete, but there were no significant differences between training methods.

As a first step towards answering this research question, a one-way mixed ANOVA was run to determine if there were any changes on each of the five metrics from pre- to post-test for either group. Specifically, the independent variable was the training group and the dependent variables were each of the metrics. Prior to conducting the one-way mixed ANOVA, five outliers were removed from the dataset. The results revealed a statistically significant difference for Path Length (F(1,7) = 11.848, p = 0.014) and Time to Complete (F(1,7) = 12.649, p = 0.012) for the robotic group. For the manikin training group, the results revealed a statistically significant difference for Path Length (F(1,7) = 10.442, p = 0.014) and Time to Complete (F(1,7) = 11.052, p = 0.013). The full results of the ANOVA and the mean values for each metric are shown in Table 6-2. These results indicate that, regardless of training group, participants decreased their time required to complete the procedure and decrease their overall path length.
Since the previous analysis looked at differences for each training group separately, a two-way mixed ANOVA was conducted to determine if there a difference in these metrics between the two conditions. Specifically, the independent variable was training group and the dependent variable were each of the metrics. There was homogeneity of variances (p > .05). Once the assumptions were verified, a two-way mixed ANOVA was computed. The results failed to reveal a significant interaction effect. see Table 6-3. Because of this, the results were examined for main effects. Specifically, the results showed a significant main effect between the pre- and post-test for time to complete (p = 0.012, F(1,15) = 8.183, η2 = .353) and path length (p = 0.027, F(1,15) = 6.019, η2 = .286), see Table 6-4 for results. This indicates that, when the groups were combined, there was a statistically significant difference for path length and time to complete over the course of training. However, the main effect of training group was found not to be significant as indicated in Table 6-5.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Robotic Group</th>
<th>Manikin Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>p-value</td>
</tr>
<tr>
<td><strong>Angle</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velocity (m/s)</td>
<td>53.63</td>
<td>0.823</td>
</tr>
<tr>
<td></td>
<td>0.42</td>
<td>0.734</td>
</tr>
<tr>
<td></td>
<td>1651</td>
<td>0.416</td>
</tr>
<tr>
<td><strong>Time to Complete</strong></td>
<td>72.3</td>
<td>0.012*</td>
</tr>
<tr>
<td>(sec)</td>
<td>19.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>244.5</td>
<td>0.014*</td>
</tr>
<tr>
<td><strong>Path Length</strong></td>
<td>244.5</td>
<td></td>
</tr>
<tr>
<td>(cm)</td>
<td>60.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.848</td>
<td></td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates significance at p < .05.
Table 6-3: Values of Interaction Effect from 2-way Mixed ANOVA

<table>
<thead>
<tr>
<th>Interaction Effect</th>
<th>Metric</th>
<th>df</th>
<th>F</th>
<th>p-value</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>0.42</td>
<td>1</td>
<td>0.53</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Velocity</td>
<td>2.95</td>
<td>1</td>
<td>0.11</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>Jerk</td>
<td>2.89</td>
<td>1</td>
<td>0.11</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>time</td>
<td>0.23</td>
<td>1</td>
<td>0.64</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Path Length</td>
<td>0.10</td>
<td>1</td>
<td>0.76</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No values were significant.

Table 6-4: Values of Main Effect for Test (Pre and Post) from 2-way Mixed ANOVA

<table>
<thead>
<tr>
<th>Test</th>
<th>Metric</th>
<th>df</th>
<th>F</th>
<th>p-value</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>0.726</td>
<td>1</td>
<td>0.408</td>
<td>0.046</td>
<td></td>
</tr>
<tr>
<td>Velocity</td>
<td>3.246</td>
<td>1</td>
<td>0.092</td>
<td>0.178</td>
<td></td>
</tr>
<tr>
<td>Jerk</td>
<td>0.916</td>
<td>1</td>
<td>0.354</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>8.183</td>
<td>1</td>
<td>**0.012</td>
<td>0.353</td>
<td></td>
</tr>
<tr>
<td>Path Length</td>
<td>6.019</td>
<td>1</td>
<td>*0.027</td>
<td>0.286</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* indicates significance at p<.05, ** indicates significance at p <.01
These results confirmed the hypothesis that objective measures of skill gains would increase with training and the skill gains for the robotic training group would be greater than or equal to the manikin training group. This finding supports the results from Chapter 3 that indicated the DHRT was as effective of a method for teaching procedural CVC skills as manikin training. These results provide support for the validation of the DHRT system and Learning Interface as an effective method for improving CVC procedural skills through objective, personalized feedback.

RQ3: Is there a relationship between objective and subjective measures of performance?

The final research question in this chapter sought to identify if there were any correlations between performance on the pre- and post- test as measured by the IJ CVC skills checklist and performance on any of the 5-objective metrics measured through motion tracking: total path length of the needle tip, time to complete the CVC insertion, average angle of insertion, velocity, and jerk. It was hypothesized that there would be a correlation between performance on the IJ CVC skills checklist and at least one of the objective measures of performance because prior research has indicated that both skills
checklists\textsuperscript{71} and motion tracking\textsuperscript{112} can be used to identify improvements in surgical skill performance. The results showed a weak, but statistically significant correlation between objective and subjective measures of performance.

In order to answer this research question, the average pre-test score was calculated by finding the average of the ratings for each question. This resulted in a value between 1 (passed all of the checklist items) and 0 (passed none of the checklist items). An example is shown in table 6-6.

Table 6-6: Example Scoring for A Pre- and Post- Test of One Participant From Each Training Group

<table>
<thead>
<tr>
<th>Test Item</th>
<th>ANIA07 - Robotic</th>
<th>ARAN07 - Manikin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selecting the appropriate site for venipuncture</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Selecting the correct ultrasound probe and using it in the appropriate orientation</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Correctly distinguishing artery and vein demonstrates compressibility of vein</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Inserting the needle at a 35-45 degree angle from the skin</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Confirming vessel entry by aspiration of blood</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of Attempts</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Final Rating (1 demonstrates competence, 0 requires further practice)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Average Score (does not include number of attempts)</td>
<td>0.80</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Next, a Spearman’s Rank Correlation was run to assess the relationship between the average pre- or post- test score as measured by the IJ CVC skills checklist and pre- or
post-test performance on the five-objective metrics measured through motion tracking: total path length of the needle tip, time to complete the CVC insertion, average angle of insertion, velocity, and jerk. The assumption of a monotonic relationship was confirmed through a visual evaluation of a scatter plot. The results showed a weak, but significant positive correlation between average pre-test score on the IJ CVC and total time to complete the pre-test \( r_s(24) = 0.46, (p = 0.021) \). These results refute the hypothesis that there would not be a correlation between performance on the IJ CVC skills checklist and any of the objective measures of performance. Prior research has indicated that skills checklists may not effectively evaluate skills needed for CVC procedures\(^7^4\) but that motion tracking\(^1^1^2\) can be used to identify improvements in surgical skill performance. These results suggest that skills checklists may be effective for evaluating CVC skills performance.

### 6.3 Discussion

The goal of this chapter was to compare two different methods of assessing CVC skill competency. This is important because most evaluation methods for CVC training utilize some form of a skills checklist\(^6^8,6^9\) which may not be sufficient feedback for developing expertise in CVC insertion procedures\(^7^5\). The main findings were as follows: (1) There were no statistically significant differences between binary skills on the IJ CVC skills checklist between the pre- and post-test or between training groups, but there was a moderately statistically significant decrease for the number of insertions from pre- to post-test; (2) There were no statistically significant changes in performance from pre- to post-test between training groups for any of the 5-objective metrics but there was a significant decrease for time and path length on the post-test. (3) There are no strong correlations between performance on the IJ CVC skills checklist and the 5-objective metrics.

Importantly, these results indicate a lack of significant difference from pre- to post-test on the IJ CVC skills checklist. This is alarming considering none of the participants had completed any training prior to the pre-test and were still evaluated as having passing skills. This finding may be attributed to the fact that prior research has indicated that checklists may not be granular enough to detect changes in performance\(^1^2^1\). Some of our observational results point to this – with many errors such as inserting the
needle in the wrong direction or arterial puncture not being classified on the IJ CVC checklist. This is problematic because the literature has indicated that most often, skills evaluation for CVC training utilizes some form of a skills checklist\textsuperscript{68,69}. However, it is encouraging to see that the number of sticks between the pre- and post-test were significantly different, regardless of the training condition as complications rates are 6 times higher when there are more than 3 insertion attempts\textsuperscript{38}.

In addition, while the results failed to detect a difference in changes in the 5-objective measures between the manikin and robotic training groups, they did identify that there were improvements on the time to complete the procedure and the path-length followed. Improvements in these metrics is important because prior work has shown that time to complete a procedure is indicative of expertise in basic surgical skills\textsuperscript{103-105,107} and the results of Chapter 4 showed that time to complete was significantly faster for experts as compared to novices. In addition, prior research has shown that overall path length of hand motions can distinguish between expert and novice performance\textsuperscript{102,105}. However, it is important to note that in Chapter 4 it was identified that a combination of path length, standard deviation of the deviations, average velocity, and time to complete could significantly predict expertise.

Finally, it is also interesting to note that there were no differences in training improvements on either subjective or objective measures between the DHRT and manikin training systems. The DHRT training system may have an advantage over standard Manikin training as it provides an objective measure of skill performance and can provide objective, consistent feedback on performance. These discrepancies between objective and subjective measures provide evidence that it may be important to distinguish between evaluation tools and training tools\textsuperscript{76}. As discussed in Chapter 5, this feedback may enable residents to better assess their own skill performance. Additionally, participants in the robotic training group were able to transfer their skills from the DHRT simulator to the manikin. Research in the field of motor skill gains indicates that participants who were tested on the same scenario they were trained in, outperformed participants who were trained in a different scenario\textsuperscript{53}. Since the participants trained in the robotic group were testing in a different scenario than they were trained in, but performed just as well as the manikin group, this suggests that the feedback provided to them is at least as effective, if not more effective, than the feedback provided to the manikin participants.
As mentioned in Chapter 5, there were several limitations to this study. Most notably, the pre- and post-tests were conducted on the same manikin that was used for the manikin training group. This limits the conclusions that can be drawn about the transference of skills. Specifically, while the DHRT training group demonstrated an ability to transfer skills to a novel situation, this type of transference is unknown for the manikin group. This is important because research has indicated that repeatedly practicing a skill in the same situation is not as effective for skill transfer as practicing the skill in variations of the scenario\textsuperscript{53}. Because the ultimate goal of this research is to develop a training system that can reduce complications associated with CVC needle insertions when transferring learned skills to clinical patients, this transfer of skill is very important. To address these issues, a long-term study is currently underway at HMC to track complications associated with CVC placement for participants trained in both groups in the surgical suite. Additionally, not all participants were able to complete the study. Because of this, it may make it difficult to draw definitive conclusions about the effectiveness of either training method. To address this, further research should focus on repeating the study to determine if the findings reported here are replicable.
Chapter 7
Conclusions and Future Work

The need for an effective training method to help reduce complications during Internal Jugular Central Venous Catheterization (IJ CVC) is undisputed. This thesis was developed to investigate the possibility of using a Dynamic Haptic Robotic Simulator to more effectively train surgical residents on the placement of CVCs by exposing them to variations in patient anatomy. Chapter 1 provided a review the relevant background literature while Chapter 2 described the process followed to design a user interface and learning feedback system for the DHRT system. The study discussed in Chapter 3 evaluated the effectiveness of the DHRT as a virtual reality trainer for increasing CVC skills and self-efficacy on CVC procedures. The study presented in Chapter 4 investigated the differences between expert and novice performance when using the DHRT system. Finally, Chapters 5 and 6 compared the effectiveness of the DHRT learning feedback system to standard manikin training when introduced as part of the surgical residency program and Hershey Medical Center. Specifically, Chapter 5 focused on the differences in self-efficacy and self-ratings over the course of training while Chapter 6 focused on the differences in both objectively and subjectively measured skill gains between pre- and post-training.

The findings of this thesis provide empirical support for the use of the DHRT system as an effective method for increasing self-efficacy and procedural skills for IJ CVC placement by exposing residents to variations in anatomy during practice needle insertions and providing consistent, objective feedback. This is important because current training methods do not expose novices to variations in anatomy prior to performing the procedure in clinical scenarios even though most complications in CVC placement can be attributed to variations in anatomy. The findings of this thesis provide the following contributions to research in this area:

1. This research identified a methodology for systematically collecting and analyzing verbal and written feedback provided to novices during surgical skills training sessions. The results can be applied in other areas of surgical training to identify and quantify the
content and quality of feedback being provided during skills training. This is important because appropriate feedback is crucial to the development of expertise in surgery.  

2. This research provided an empirical basis for the development of a graphical user interface and learning feedback system for the Dynamic Haptic Robotic Training System for ultrasound-guided internal jugular catheterization based on verbal and written feedback typically provided during a skills training session on this procedure. Specifically, the second study in Chapter 2 provides an example of how to empirically test the usability of such an interface and incorporate it into a stand-alone system for training surgical residents. These results can be used as a basis for developing learning feedback systems in virtual reality training devices for surgical procedures and contribute to the rapidly expanding field of virtual reality in surgical training.  

3. This research provides an empirical basis for the validation of a virtual reality haptic training system for surgical procedures when paired with observer feedback. Specifically, the results of the study conducted in Chapters 3 demonstrated that a virtual reality haptic training device increased skill performance and self-efficacy as effectively as a manikin simulator. However, manikin simulators have been shown to only temporarily improve procedural skills. These results contribute to the growing body of literature that supports the use of virtual reality haptic simulators as an effective method for training surgical skills.  

4. This research contributes to the knowledge base of skill evaluation for Internal Jugular Central Venous Catheter placement. Specifically, this research empirically and methodically identified 4 objective metrics for evaluating surgical skill proficiency in placing an Internal Jugular Central Venous Catheter: average velocity during insertion, total path length of the needle tip during insertion, total time to complete the procedure, and standard deviation of the deviations. These results can be used to drive the development of objective methods for evaluating surgical skill performance rather than just checklists which may not provide helpful feedback.  

5. This research provided empirical evidence that showed that the Dynamic Haptic Robotic Trainer and Personalized Learning Interface System is an effective method for improving self-efficacy on CVC procedures. Specifically, the studies conducted in this
thesis demonstrate that completing training on the DHRT system resulted in increased confidence of ability to complete the procedure which provides support that virtual reality can be an effective method for improving resident comfort with surgical procedures. These results can be used to support and encourage the development of stand-alone training systems for improving novice comfort with performing surgical procedures. This is important because a lack of self-efficacy is related to poor skill performance.

6. This research provided evidence that suggests surgical residents may not have a strong ability to accurately self-assess their performance. Additionally, the findings of the study presented in Chapters 5 and 6 indicate that providing residents with consistent, specific, and objective feedback can help them more accurately assess their performance. These results can be used to develop feedback methods that help novices in surgery accurately assess their performance and therefore reduce overconfidence.

7. This research provided empirical evidence that showed that the Dynamic Haptic Robotic Trainer and Learning Interface System was an effective method for improving objectively measured skills used in CVC procedures by providing direct, specific feedback. Specifically the studies conducted in this thesis demonstrate that completing training on the DHRT resulted in improved CVC skills. This provides support that virtual reality trainers can be an effective tool for training and providing feedback to surgical residents. These results can be used to support the development of virtual reality simulators to train novices on skills for surgical procedures and then evaluate their competency on these skills. This is important because there is currently no standardized method of effectively and objectively evaluating CVC skills.

Limitations and Future Work

The results of the studies presented in this thesis provide empirical evidence that the DHRT system is at least as effective as standard manikin training for increasing CVC self-efficacy and procedural CVC skills. However, all of the studies presented in this thesis only examine skill performance on a manikin trainer. While these studies provide evidence that residents trained on the DHRT system can transfer their skills to a novel situation (the manikin simulator), it remains unknown if residents trained on the manikin
simulator can transfer their skills to a novel situation. Therefore, future research should examine the changes in self-efficacy and skill performance when transferred to a situation that is novel for both the DHRT and manikin training conditions.

Additionally, the results presented in Chapter 5 of this thesis revealed that residents may be over-confident in their ability to effectively complete CVC procedures. The results presented in Chapter 6 indicate that current evaluation methods for CVC skills placement, using checklists filled out by an observer, may not be sufficient to determine surgical skill competency. Future research should continue to investigate the relationship between self-efficacy and skill performance as well as rigorously compare self-ratings of performance and observer-ratings of performance to objective measures of performance.

Finally, and most importantly, the conclusions drawn from the results of the research presented in this thesis are limited to CVC performance on a manikin simulator. Ultimately, the goal of this research is to contribute to the development of a training system that will reduce complications associated with CVC placement on actual patients. Therefore, future research should examine how utilizing the DHRT as a training method impacts complications in clinical scenarios.
References


70. Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. The American journal of surgery 2007;193:551-5.


Appendix A

Patient Profiles

Vessel Placement for Initial 8 Patient Profiles

**Patient Profiles:**

**Scenario 1:**
41 year old Hispanic female with end stage kidney disease presents with severe hyperkalemia requiring dialysis.
Factors to include:
- thicker skin

**Scenario 2:**
66 year old male is admitted to the ICU with a closed head injury after a motor vehicle collision and requires a central line for 3% Saline administration.
Factors to include:
- thin skin
**Scenario 3:**
25 year old white male presents after sustaining a gunshot wound to the abdomen. He lost a significant amount of blood at the scene and has been tachycardic and hypovolemic since his arrival in the trauma bay.
Factors to include:
- constricted vessels

**Scenario 4:**
62 year old African American male presents after falling from roof. He has multiple injuries and is hemodynamically unstable.
Factors to include:
- thicker skin
- constricted vessels
**Scenario 5:**
15 year old white female presents after ingesting a large quantity of unknown prescription medications, with somnolence and severe electrolyte abnormalities.
Factors to include:
- thin adipose tissue
- anatomical variant?

**Scenario 6:**
82 year old white female presented to the ED severely dehydrated and was found to have a bowel obstruction. She is hemodynamically unstable and requires a central line.
Factors to include in this scenario:
- thin skin
- small amount of adipose tissue
- constricted vessels
Scenario 7:
33 year old morbidly obese male with DM, COPD, and CHF presents with necrotizing fasciitis and requires a central line prior to operative debridement for hemodynamic support.
Factors to include in this scenario:
- thick skin
- large amount of adipose tissue

Scenario 8:
24 year old male presented as level I trauma activation after his vehicle collided with oncoming traffic. His injuries include a left open femur fracture, for which he requires operative management. He is becoming increasingly tachycardic and hypotensive and the decision is made to place a central line.
Factors to include in this scenario:
- thicker (but not extremely thick skin)
- medium adipose tissue
- constricted vessels
<table>
<thead>
<tr>
<th>Scenario 1:</th>
<th>66 year old male is admitted to the ICU with a closed head injury after a motor vehicle collision and requires a central line for 3% Saline administration. Height: 5’9” Weight: 185.1lbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 2:</td>
<td>15 year old white female presents after ingesting a large quantity of unknown prescription medications, with somnolence and severe electrolyte abnormalities. Height: 5’4” Weight: 131.1lbs</td>
</tr>
<tr>
<td>Scenario 3:</td>
<td>39 year old Female requires IVF. However, after several attempts by the floor nurses, the ICU nurses, and residents, you cannot obtain peripheral access. The decision is made to place a central line. -standard thin patient</td>
</tr>
<tr>
<td>Scenario 4:</td>
<td>A 28 year old M with Guillain Barre is in need of plasmapheresis and requires a central line. -standard thin patient</td>
</tr>
<tr>
<td>Scenario 5:</td>
<td>A previously healthy, thin, 25 year old M presents with a closed head injury after a skate-boarding accident. A central line is needed for hypertonic saline and possible vasopressors. -standard thin patient</td>
</tr>
</tbody>
</table>
**Scenario 6:**
25 year old white male presents after sustaining a gunshot wound to the abdomen. He lost a significant amount of blood at the scene and has been tachycardic and hypovolemic since his arrival in the trauma bay.
Height: 5'10”
Weight: 176.5lbs

**Scenario 7:**
24 year old male presented as level I trauma activation after his vehicle collided with oncoming traffic. His injuries include a left open femur fracture, for which he requires operative management. He is becoming increasingly tachycardic and hypotensive and the decision is made to place a central line. Muscular neck.
Height: 5’11”
Weight: 206.6lbs
-more contracted vessels

**Scenario 8:**
44 year old African American male with minimal change disease is in need of dialysis. He has many scars from prior central access sites.
-thickened skin
-scars overlying puncture site – tougher to puncture through

**Scenario 9:**
41 year old Hispanic female with end stage kidney disease presents with severe hyperkalemia requiring dialysis.
Height: 5’2”
Weight: 158.9lbs
-tougher skin
-more adipose tissue
<table>
<thead>
<tr>
<th>Scenario 10:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 62 year old African American male presents after falling from roof. He has multiple injuries and is hemodynamically unstable. Height: 5’11” Weight 188lbs -tougher skin -contracted vessels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 11:</th>
</tr>
</thead>
<tbody>
<tr>
<td>An 82 year old white female presented to the ED severely dehydrated and was found to have a bowel obstruction. She is hemodynamically unstable and requires a central line. Height: 5’3” Weight: 140.0lbs -thin skin, thin adipose tissue -superficial vessels -contracted vessels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 12:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 33 year old morbidly obese male with DM, COPD, and CHF presents with necrotizing fasciitis and requires a central line prior to operative debridement for hemodynamic support. Height: 5’7” Weight: 282.2lbs -large amount of adipose tissue -difficult to see vessels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 13:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 52 year old obese female with atrial fibrillation, DM, and peripheral vascular disease presents with massive hematochezia and requires a central line for blood products. -large amount of adipose tissue -contracted vessels</td>
</tr>
</tbody>
</table>
Scenario 14:
65 year old obese African American Female was found down by her boyfriend. In the ED, they have difficulties obtaining stable IV access and decide she requires a central line.
-thick skin
-large amount of adipose tissue

Scenario 15:
67 year old M has been in the ICU for weeks after sustaining injuries in a motor vehicle crash. He is edematous and recently developed acute renal failure. He requires central venous access for initiation of acute hemodialysis.
-swollen skin and subcutaneous tissue
-difficult to see vessels

Scenario 16:
A 65 year old morbidly obese M is intubated and in the ICU with CHF exacerbation. He requires a central line for monitoring of CVP and administration of medications. He has a very thick neck that obscures anatomical landmarks and appears extremely edematous.
-swollen skin and subcutaneous tissues
-large amount of adipose tissue
-difficult to see vessels

Practice/Reference Scenario:
Scenario 17:
A 29 year old male in the ICU for a head injury after an ATV accident.
Height: 5' 8”
Weight: 155lbs
Appendix B

Surveys and Checklists

Central Line Self Efficacy Survey

CVC SELF-EFFICACY SURVEY

PARTICIPANT BACKGROUND

Participant ID (this unanimous identifier will be used in subsequent surveys):

<table>
<thead>
<tr>
<th>Last two characters of your first name (e.g. Sally would be LY)</th>
<th>2 digits of the numerical day of your birth (e.g. born Dec 4th would be 04)</th>
<th>Last two characters of the city in which you were born (e.g. Born in Atlanta would be TA)</th>
</tr>
</thead>
</table>

Your age? ____________

Your gender? □ Female □ Male

PGY Level? (Lab residents please count each year in the lab as 1. For example a first year lab resident would be a PGY3) __________________________________________________________________________

How many times have you completed the following (please estimate to the best of your ability):

___ I have not had ultra-sound guided training
___ Had ultrasound-guided CVC training on a mannequin (simulator)
___ Performed CVC procedures on a patient under supervision

Please list any medical training or experience you have had outside of Medical School
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
2. **PLEASE RATE YOUR CONFIDENCE IN THE FOLLOWING CVC INSERTION PROCEDURES**

Please rate your *level of confidence* in the following skills:

<table>
<thead>
<tr>
<th>Skill</th>
<th>Not at all confident</th>
<th>Extremely confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obtaining a clear image of the target vessel using the ultrasound machine</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Locating the needle on the ultrasound image</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Identifying the correct site of insertion on the skin for the introducer needle</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Using <em>tactile feedback</em> (sense of feel) during placement of the line</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Using the <em>ultrasound image/transducer</em>, identifying the correct vessel for puncture</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Using <em>tactile feedback</em>, identifying the correct vessel for puncture</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Advancing the introducer needle slowly and steadily</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Modifying the needle trajectory based on <em>ultrasound feedback</em></td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Identifying when the introducer needle is in the correct location</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Using <em>tactile feedback</em> to help guide the introducer needle</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein in one attempt</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein in multiple attempts</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Conducting the entire procedure without any mistakes</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>Conducting the entire procedure on a CVC simulator (mannequin)</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
</tbody>
</table>

*Please provide rationale for your ratings*
# Training Evaluation Surveys

**PLEASE RATE YOUR LEVEL OF AGREEMENT ON THE MANNEQUIN TRAINING YOU RECEIVED TODAY**

<table>
<thead>
<tr>
<th>The Mannequin Training:</th>
<th>Completely Disagree</th>
<th>Completely Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>... made me sensitive to how patient anatomy impacts the force required to advance needles in the human body</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... made me sensitive to patient anatomies' impact on vessel location in the body</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me understand how to <em>modify</em> CVC insertion procedures based on patient anatomy</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me learn where to insert the introducer needle during the procedure</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me learn to distinguish the vein from the artery on the ultrasound image</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me learn how and when to modify my needle's trajectory during the insertion</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me correctly identify when I had successfully inserted the needle</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... helped me understand how to use <em>tactile feedback</em> (sense of feel) during placement of the line</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
<tr>
<td>... was an effective method for learning the CVC process</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td></td>
</tr>
</tbody>
</table>

The *Patient Profiles* made me think differently about how variations in patient anatomy impacts CVC placement

<table>
<thead>
<tr>
<th>Please provide rationale for your ratings <em>please do not leave blank</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

## PLEASE RATE YOUR LEVEL OF AGREEMENT ON THE HAPTIC ROBOTIC TRAINING YOU RECEIVED TODAY

<table>
<thead>
<tr>
<th>The Haptic Robotic Training:</th>
<th>Completely Disagree</th>
<th>Completely Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>… made me sensitive to how patient anatomy impacts the force required to advance needles in the human body</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… made me sensitive to patient anatomies’ impact on vessel location in the body</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… helped me understand how to <em>modify</em> CVC insertion procedures based on patient anatomy</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… helped me learn where to insert the introducer needle during the procedure</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… helped me learn to distinguish the vein from the artery on the ultrasound image</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… helped me learn how and when to modify my needle’s trajectory during the insertion</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>…helped me correctly identify when I had successfully inserted the needle</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>…helped me understand how to use <em>tactile feedback</em> (sense of feel) during placement of the line</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
<tr>
<td>… was an effective method for learning the CVC process</td>
<td>□ □ □ □ □</td>
<td></td>
</tr>
</tbody>
</table>

Please provide rationale for your ratings *please do not leave blank*
# IJ CVC Skills Checklist

<table>
<thead>
<tr>
<th>Task</th>
<th>Satisfactory</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbalizes consent, Universal Precautions, and Time Out</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Prepares catheter kit: flushes catheter, ensures proper equipment is</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintains sterile technique</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Selects appropriate site for venipuncture</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Simulates injecting local anesthesia into skin at appropriate site</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Selects correct ultrasound probe and applies in appropriate</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>orientation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correctly distinguishes artery and vein: demonstrates</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>compressibility of vein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inserts introducer needle at 35-45° angle from the skin</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Venipuncture successful</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Number of attempts: ____</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirms vessel entry by aspiration of blood (blue)</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Removes syringe while occluding the hub of needle</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Inserts wire into needle and advances without resistance</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Maintains control of the wire</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Removes introducer needle</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verbalizes skin incision with #11 blade</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verbalizes dilator insertion and removal</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Passes catheter into the vessel and removes wire</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verbalizes correct distance for insertion of catheter (testers may</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>prompt for distance)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulates aspirating blood through catheter</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verbalizes securing the catheter with suture and applying dressing</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verbalizes placing order for CXR</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economy of time and motion</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Many unnecessary/disorganized movements</td>
<td>Maximum economy of movement and efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final Rating</th>
<th>Other Summative Comments:</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Demonstrates competence</td>
<td></td>
</tr>
<tr>
<td>☐ Requires further practice</td>
<td></td>
</tr>
</tbody>
</table>
Post-Test Checklist

PARTICIPANT CODE ____   ____   ____   ____   ____   ____
Time to complete procedure:______________________

Please identify how successful the resident was in performing each of the following skills without any assistance or prompting:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Satisfactory</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selecting the appropriate site for venipuncture</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Selecting the correct ultrasound probe and using it in the appropriate orientation</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Obtaining a clear image of the target vessels using the ultrasound machine</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Correctly distinguishing artery and vein: demonstrates compressibility of vein</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Inserting the needle at a 35-45 degree angle from the skin</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Locating the needles position on the ultrasound image</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Advancing the introducer needle slowly and steadily</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Placing the introducer needle at the center of the vein</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Number of attempts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirming vessel entry by aspiration of blood</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Conducting the entire procedure without any mistakes</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

Please identify if any of the following occurred during the procedures (check all that apply and provide details on next page):

____ Inadvertent arterial puncture
____ Unsuccessful insertion
____ Assistance on procedure
____ Other ________________________________

Please provide rationale or notes about your ratings or complications in the procedure, including any of the items checked in the previous question.

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
Appendix C

Training Scripts

Manikin Training Script

[Read this page 1 time before training each individual]

[ASK FOR PARTICIPANT ID]

[ENSURE PARTICIPANT IS WEARING GLOVES and Dots]

[VERBAL] You will be given different patient profiles and asked to insert a guide needle for each patient. When prompted, begin the procedure with using the ultrasound to find your target location and when you are confident in the placement of your needle for inserting the guide-wire, set down the ultrasound probe. You may assume that your patient has already been draped and consent obtained.

Throughout the entire procedure please verbalize your actions and thoughts, including what you are doing or looking for and why.

Tell me:

When and how you have identified the vein

When you have identified the needle in the ultrasound

When the needle is in the center of the vein
Using the Robot Script

[read once for each participant]

You will be using this robot to simulate a needle insertion for a central venous catheter placement. I will show you a demonstration and then you will be able to perform a practice trial.

Here is the ultrasound probe [indicate probe], your ultrasound screen, and your syringe [indicate syringe]. At the start of each trial, the ultrasound probe needs to be placed in its holder on the left. [Place probe] The needle needs to be placed in the holder on the right. [Place needle]. I will verify their placement between each trial.

To start the simulation, I’ll hit the run button. [Start Program] A black box will appear and after approximately 15 seconds an ultrasound image will appear. At this point you will be ready to start the test.

This is the scanning surface. The anatomical landmarks are labeled for the area around the right internal jugular vein. This is the patient’s cheek, clavicle, and sternocleidomastoid. [point]

Although the system is simulated you can move the probe around in the same way you would on the mannequin [demonstrate], however please be aware that if you move too far away from the testing area [demonstrate] you will end the trial early. Also observe that you can pull back on the plunger to simulate aspiration [demonstrate].

First, you will use the probe to scan the area and find the two vessels. Just like a real ultrasound, I can press down with the probe to can check for deformation in the vein. [press down] The artery on the left is pulsatile. You can see that as I rotate the probe the screen moves with it. [rotate probe]

You can see I am in the correct position so now I will try to insert the needle at a 45 degree angle at the location of the probe. [begin insertion]
You can also see on the screen several lines. These represent tissue deformation around the site. You can see the tissue deformation and now you can see the needle has appeared in the ultrasound image where it has crossed the plane.

When inserting the needle, the robot will provide a variety of force feedback similar to which you would feel working on the varying anatomies of patients.

An important task when inserting is to make sure you aspirate the needle; if you hit a vessel while aspirating you will see a red bar appear across the screen if I am in the artery, and a blue bar will appear if I am in the vein. This indicates flash in the syringe.

When you think you are in the center of the vein, inform me and hold position until I tell you the test is over.

I will reset the system between each trial.

Do you have any questions? [pause]

[Trial run - use separate script]
Robot Training Script

[Read this page 1 time before training each individual]

[VERBAL] You will be given different patient profiles and asked to insert a guide needle for each patient. When prompted, begin the procedure with using the ultrasound to find your target location and when you are confident in the placement of your needle for inserting the guide-wire, set down the ultrasound probe.

Throughout the entire procedure please verbalize your actions and thoughts, including what you are doing or looking for and why.

Tell me:

When and how you have identified the vein

When you have identified the needle in the ultrasound

When the needle is in the center of the vein
Appendix D

Coding Handbook for Feedback on Central Line Placement

Handbook for

Coding in NVivo:

Hershey Medical

Central Line Insertion Training Feedback
General Coding Methods:

- Coding will be conducted only during the discussion between the medical student and the instructor (Katelyn); read but ignore the students’ narrations of their attempts to insert CLI
- To ensure that the same passages are highlighted, all coding will be done on a sentence-by-sentence basis. That is, despite how short a reference to a node may be, the entire sentence in which it is included will be highlighted.
- A sentence is ended exclusively by a period (.), or question mark (?). Anything separated by any other punctuation mark such as commas, semicolons, colons, or ellipses (…) should still be considered as one sentence.
- More than one node can exist in any given sentence.
- Some nodes may be discussed multiple times. Only code the first time a certain issue is mentioned.
- A node will be discussed for more than one sentence, choose the sentence that you find most directly discusses the particular node
- Child nodes will often be determined by context and tend to be added in with most nodes such as a praise or mistake
- The following will be a guide for identifying instances to code each node:

Selecting the appropriate site for venipuncture:

- Centered on where needle is inserted at the skin, not the actual vein.
- Code for any time direction of needle (longitudinal/latitudinal), at site of insertion
- Any time placement of insertion is mentioned: could be body, or anatomical phrasing (e.g. “want to be more medial”)

Selecting the correct ultrasound probe and using it in the appropriate orientation:

- Having good orientation is directly mentioned
- The discussion of correctly checking orientation (with finger), or figuring out probe is mentioned

Obtaining a clear image of the target vessels and using the ultrasound machine:

- Any instance in which Katelyn directly mentions the images of the vessels on the ultrasound
- Typically, this node can be found when Katelyn discusses how an ultrasound works, and projects images in perpendicular lines. Any point Katelyn discusses this with the context of the vessels, and seeing their cross-section would be appropriate for coding

Correctly distinguishing artery and vein:
- Katelyn directly mentions compressibility of a vein when commenting on identifying the vessel
- Child nodes:
  o Correctly distinguishing with compressibility
  o Correctly distinguishing with vein without compressibility
    ▪ Anatomically
    ▪ Pulsatility

**Inserting the needle at a 35-45 degree angle from the skin:**

- Any time the angle of the insertion is mentioned
- Words like “steep” or “shallow” may be used to describe the angle of insertion

**Locating the needle’s position on the ultrasound image:**

- Typically, this node can be found when Katelyn discusses how an ultrasound works, and projects images in perpendicular lines.
- The first time the needle is mentioned with this explanation (before, during, or immediately after) is the sentence to code

**Advancing the introducer needle slowly and steadily:**

- Anytime Katelyn mentions going in slowly and steadily
- Anytime Katelyn discusses the torqueing of the needle during the insertion

**Placing the introducer needle at the center of the vein:**

- Any instance Katelyn discusses the needle’s position in relation to the center of the vein (e.g. going through and through,

**Confirming vessel entry by aspiration of blood:**

- Anytime aspirating is mentioned in the context of confirming vessel entry
- Any instance in which aspirating is discussed in the context of getting flash back, or knowing whether or not student went “through and through”
- Any time where Katelyn discusses “aspirating easily”
- May often overlap with **Aspirating the Entire Time**

**Conducting the entire procedure without any mistakes:**

- Instances in which Katelyn says she has no critiques at the beginning of her feedback
- If Katelyn almost immediately asks for questions or discussing the rest of the procedure

**Aspirating the entire time:**
- Aspirating the entire time is directly mentioned
- Any time Katelyn discusses aspirating as soon as needle is inserted in body

**Hand position on the syringe (how they're holding it):**

- Usually can be identified when Katelyn says “holding like this” or a close derivative
- Any time a student had to “walk hands up” the syringe
- Likely discusses how awkward the positioning is in the context

**Anchoring hand on body to ensure stability (need to get exact name):**

- Any time Katelyn mentions not wanting free-floating hands during the procedure
- Typically, the phrase “brace yourself” is used
Appendix E

Self-Evaluation Training Book

Central Venous Catheter Training

*Summer 2016*

*Manikin*

<table>
<thead>
<tr>
<th>Participant ID #</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Last two characters of Mother's First Name</th>
<th>Last two characters of Birth City</th>
<th>Two digits of Birth Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e.g. Sally would be LY)</td>
<td>(e.g. Atlanta would be TA)</td>
<td>(e.g. Jan would be 01)</td>
</tr>
</tbody>
</table>
Instructions

This book is provided to log your practice attempts as you complete your CVC skills training. For the first part of this training, you are to complete the following:

**Needle Insertion Practice**

- Complete 22 needle insertion practices which are focused on correctly inserting the needle into the center of the vein
- Complete the 4-question feedback form after you complete each needle insertion practice

**Each practice attempt is considered complete when you:**

- Are satisfied with your placement of the needle in the vein
  OR
- Puncture the artery
Needle Insertion #1

Date Completed:

Complete the needle insertion before flipping page

Answer the following questions before continuing to the next trial:

Rate your performance

<table>
<thead>
<tr>
<th>Very Poor</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Great</th>
</tr>
</thead>
</table>

Number of attempts* to access vessel

List any errors you made:

Did anyone give you feedback?

YES ☐ NO ☐

* Attempts are defined as a large manipulation with the needle or withdrawing the needle and re-entering