Spring 2014

The creation and validation of an augmented reality orthopaedic drilling simulator for surgical training

Brian Douglas Johns
University of Iowa

Recommended Citation

Johns, Brian Douglas. "The creation and validation of an augmented reality orthopaedic drilling simulator for surgical training." PhD (Doctor of Philosophy) thesis, University of Iowa, 2014.
https://doi.org/10.17077/etd.x8t4uged

Follow this and additional works at: https://ir.uiowa.edu/etd
Part of the Industrial Engineering Commons

THE CREATION AND VALIDATION OF AN AUGMENTED REALITY ORTHOPAEDIC DRILLING SIMULATOR FOR SURGICAL TRAINING

by

Brian Douglas Johns

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Industrial Engineering
in the Graduate College of
The University of Iowa

May 2014

Thesis Supervisor: Associate Professor Geb W. Thomas

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Brian Douglas Johns

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Industrial Engineering at the May 2014 graduation.

Thesis Committee:

Geb W. Thomas, Thesis Supervisor

Donald D. Anderson

Matthew D. Karam

Thomas Schnell

Priyadarshini R. Pennathur

I have a theory that the truth is never told during the nine-to-five hours.
-Hunter S. Thompson

ACKNOWLEDGMENTS

The title page shows only one author, but this work would not be possible without the support and guidance from many. From the academic support to the moral support, this dissertation is the result of a team effort.

I would like to express my gratitude to my advisor, Geb W. Thomas, for his excellent guidance and mentoring through this arduous journey called graduate school. I would like to thank my secondary advisor, Donald D. Anderson, without his support this research in orthopaedics would not be possible. I would like to acknowledge the additional committee members, Dr. Matthew Karam, Thomas Schnell, and Priyadarshini Pennathur for their commitment to my work. In addition, the research trials and experiments were made possible with the help of Dr. Jenniefer Kho, Brian Westerlind, Kevin Dibbern, and Steven Long.

My family deserves the largest recognition for this work. I would not possess the fortitude to accomplish this feat without their upbringing. Additionally, I would like to thank my sister, Stacia L. Johns, and friend, Eric M. Salter, for their assistance editing this dissertation. Lastly, to all my comrades pursuing secondary degrees in engineering, thank you for the support over the countless late nights and many years together. We are all going to make it in the end.

ABSTRACT

Standard surgical repair of intertrochanteric hip fractures requires accurate placement of a wire across the fracture using static fluoroscopic images. Few practice methods exist for perfecting this wire navigation skill outside the operating room. The objective of this research is to further understand skill development for orthopaedic drilling using a validated simulator, enabling more effective instruction and training. This includes the investigation of the relationship between practice and skill acquisition in conjunction with specific differences between experts and novices.

This work details the creation and validation of an augmented reality wire navigation simulator for training orthopaedic drilling. This novel augmented reality simulator combines real-world objects, such as a surgical drill and synthetic bone, with virtually generated, radiation-free radiographic imaging. The central hypothesis is that an augmented reality wire navigation simulator will demonstrate construct validity and improve orthopaedic drilling skill through simulation training.

This work identifies the differentiation of skill between experienced surgeons and novices completing the wire navigation task, demonstrating construct validity for the developed simulator. It also demonstrates that experienced surgeons are more accurate than novices in orthopaedic drilling ($F(2, 39) = 3.721$, $p = 0.033$). This provides evidence supporting the simulator's construct validity and value as a training and assessment tool in wire navigation of the proximal femur.

Although the study was unsuccessful in providing sufficient evidence that training on the simulator directly transfers to more realistic drilling tasks, it revealed several discoveries about acquiring wire navigation skill. This work establishes a relationship between skill acquisition and practice for the wire navigation task. This learning curve shows that skill acquisition occurs much more slowly in wire navigation than previously assumed. The wire accuracy (i.e., tip-apex distance) is predicted to improve 0.1 to 0.3

millimeters with each successive practice repetition of the wire navigation task (95% CI, $p < 0.001$). In addition, the time to complete wire navigation improves between 0.4 and 2.0 seconds each subsequent practice trial (95% CI, $p < 0.001$).

The developed simulator also identified several flaws in novice technique. First, novices do not account for the anteversion of the femoral neck indicating that the inclination angle is difficult for novices to understand and accurately drill from radiographic images. Another discovered flaw of novice orthopaedic residents is their lack of ability to accurately estimate distances in radiographic images. Novices were found to incorrectly estimate the wire accuracy by an average of 12.4 millimeters. Overall, this work establishes new findings, which can be used for future simulation development and coaching, enabling safer, more effective training methods for surgical residents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 – INTRODUCTION

1.1 Overview

This research explores the creation and validation of an augmented reality wire navigation simulator for orthopaedic drilling. This study advocates supplementing the current orthopaedic apprenticeship-training model with simulator training. It describes the design and validation of an effective training simulator that integrates the primary stimulus-response cues of the wire navigation task, an important task in many orthopaedic trauma surgeries. This novel augmented reality simulator combines real-world objects, such as a surgical drill and synthetic bone, with virtually generated, radiation-free radiographic imaging.

Experiments with the simulator demonstrate construct validity, and provide evidence that the simulator is an effective training device for orthopaedic drilling. Construct validity is established by measuring differences in the performance between novices and experienced surgeons on the simulator. These differences provide evidence that supports the notion that the simulator effectively exercises the same skills as the real drilling task. Results show that experienced surgeons are more accurate than novices; however, experienced surgeons do not significantly differ in the duration or number of radiographic images used to complete the drilling. Although similar studies have hypothesized that experienced surgeons are more accurate in completing a drilling task [1, 2], this work is the first to provide statistically significant evidence showing that experienced surgeons outperform novices in drilling accuracy on a simulator.

This study also examines the transfer of training from the simulator to a drilling task using real radiographic imagery. The results were unexpected. Simulator-trained residents demonstrated little to no improvement after simulator training. Further investigation led to the discovery of previously unknown details that influence orthopaedic drilling skill, which may explain the unexpected results. A learning curve study revealed that the orthopaedic drilling skill develops slowly, requiring many repetitions over several days for observable improvement. Another study performed suggests that novices lack the skill to account for the inclination angle of the femoral neck while drilling and the ability to accurately estimate distances in the radiographic imagery. Each of these previously unidentified lacking components could limit the novices' ability to make fundamental decisions in the drilling task.

This work advances current knowledge in simulation and orthopaedic training through the creation of a novel simulator capable of differentiating the skill of experienced surgeons and novices in the wire navigation task. Furthermore, it establishes the relationship between practice and skill development for orthopaedic wire navigation of the proximal hip. This study also reveals weaknesses in novices' perception of distance and angles in radiographic images. Overall, this work establishes new findings that provide a foundation for future simulation development and coaching, while enabling safer training methods for surgical residents.

1.2 Problem Statement and Objective

In 2010, medical errors were attributed to an estimated 98,609 deaths in the United States [3], ranking errors as high as the 6th leading cause of death in the nation [4]. Many of these errors are attributed to technical errors caused by lack of skill. Technical complications ranked the third highest cause of adverse medical events, ranking behind only drug complications and infection [5].

Contrary to common belief, surgical skill is not the result of steady hands or an innate ability possessed only by the genetically gifted. Surgical skill, like any other skill, develops with extended, effortful practice and timely performance feedback. In a study reviewing 135 volume-outcome studies spanning over 27 surgical procedures, a strong correlation existed between volume (i.e., practice) and better outcomes [6]. To further identify the gap between actual and adequate training, residency program directors deemed 121 procedures essential to practice general surgery and 1,022 graduating residents were assessed on the volume of individual procedures. Thirty-one of the 121 procedures were performed less than once on average, with only 18 of the 121 procedures performed, on average, more than ten times throughout residency [7]. This brings about an interesting question: how should residency programs increase surgical competency without exceeding the mandated maximum number of working hours per week all while avoiding increased safety risks, errors, and costs? One potential solution is the implementation of validated training simulators.

One surgical discipline currently lacking in validated training simulators is orthopaedics. Orthopaedic surgery depends heavily upon the technical skill of individual surgeons and holds great potential for research improving surgical training. Hip fracture surgery is a target for simulation implementation because it is one of the most serious current problems afflicting the aging population, with incidence estimates exceeding 247,000 per year [8]. In addition, hip fractures have the highest cost of any orthopaedic procedure after surgery, averaging more than $11,000 per year in extra health care costs [9]. The most important element of the development is not the technology or physical likeness, but the validity of the simulator. Validity of a simulator is the extent which the simulation induces the same physiological responses as the real operating task. The validity of a simulator gives merit of the use of a simulator as a training tool. Of few hip fracture simulators developed, a mere two studies complete validation experiments [1, 2].

Validation is a final step after many hours of simulator development. Surgical training programs are unlikely to adopt a simulator into training curricula without evidence that simulator practice improves skill in the operating room. The small number of validated simulators is one of the determining factors limiting acceptance of simulators into medical training programs.

The objective of this research is to further understand skill development for orthopaedic drilling using a validated simulator, enabling more effective instruction and training. This includes the relationship between practice and skill acquisition in conjunction with specific differences between experts and novices. The overreaching goal is to enable effective development of training simulators through increased understanding while preserving patient safety.

1.3 Central Hypothesis and Specific Aims

The central hypothesis of this study is that an augmented reality wire navigation simulator will demonstrate construct validity and improve orthopaedic drilling skill. Establishing construct validity of a developed augmented reality wire navigation simulator will reveal specific differences between novice and expert technique. The identification of these differences leads to a unique understanding of wire navigation skill, consequently leading to more effective surgical training and simulator development. In addition, providing evidence training on the developed augmented reality simulator improves skill in a more realistic scenario demonstrates it is effective in orthopaedic training, and, by extension, actual orthopaedic surgery. This evidence creates opportunities for using augmented reality in other orthopaedic and medical simulators, facilitating an alternative to the more commonly used force-feedback devices in simulations involving haptic sensations. The investigation in wire navigation training also establishes a relationship between repetition and skill acquisition in the wire navigation task, a previously unexplored area in the orthopaedic surgical field.

For the first aim, this research demonstrates the design of an augmented reality simulator for wire navigation. This development takes a unique approach by developing a training simulator through the integration of selected stimulus-response cues of the surgical drilling rather than trying to recreate reality. The accuracy of the developed simulator's motion tracking and calibration will be assessed ensuring the simulator is accurate enough for training purposes.

The second aim addresses the simulator's construct validity by demonstrating novices and experienced surgeons perform differently on the simulator. This is established through quantifying differences between novice and experienced surgeons on the simulator. Surgeons were hypothesized to outperform novices in categories such as wire accuracy, duration, number of drilling attempts, and images used throughout the simulation trial. The results showed experienced surgeons only significantly outperformed novices in wire accuracy. This demonstrates the simulator measures one specific feature of the surgical skill that improves with experience.

The third aim addresses the skill transfer from the simulator to a more realistic wire navigation simulation task. After a short instructional session, first-year orthopaedic residents were separated into two equal groups. One group underwent four training trials on the augmented reality wire navigation simulator. The other group received no hands-on training. Both groups then performed two wire navigation assessment trials on an artificial hip model using real fluoroscopy. Although it was hypothesized the simulator-trained group would outperform the untrained group, no significant effect was discovered. Several additional experiments demonstrated the observed result might be the effect of insufficient training time or due to limitations novices' ability to interpret feedback during the task.

1.4 Previous Exploration in Surgical Training

This research is the culmination of the collaboration between the Department of Industrial Engineering and Orthopaedic Biomechanics Laboratory at the University of Iowa, working in conjunction with the Department of Orthopaedics and Rehabilitation at University of Iowa Hospitals and Clinics. In 2010, graduate students Andrew Kern and Thad Thomas, with principal investigator Professor Donald D. Anderson from the Orthopaedics Biomechanics Laboratory, developed a tibial plafond fracture simulator using artificial bone fragments inside a surrogate soft tissue model. During simulator trials, subjects used Kirschner wires with the aid of fluoroscopy for reducing the fracture. This simulator was developed for computation of pre-planning fracture reduction strategies and assessing contact stresses of reconstructed fractures using numerical solving methods.

Preliminary trials were run, and the data revealed the developed tibial plafond simulator could also be used as a training device. A second iteration of the tibial plafond fracture simulator was developed and additional trials commenced in late 2012. With the success of the tibial plafond fracture simulator, the Orthopaedics Biomechanics Laboratory recruited the help of the Industrial Engineering Department at the University of Iowa, more specifically, my advisor, Professor Geb W. Thomas, for future training simulator development. Professor Thomas has a background in the development of training simulators, ranging from breast cancer detection simulators to dental cavity detection simulators.

In early 2012, I was enrolled as part of the surgical simulation research team. At that time, fluoroscopy-based wire navigation was a potential avenue for simulator development due to the commonality, transferability, importance in success of the surgery, and the skill required to perform the task. The initial design idea emerged from a simple graphic of a potential simulator placed in a research grant proposal by Dr. Jenniefer Y. Kho, an orthopaedic resident at the University of Iowa, where the objective was to position a guide wire while direct view of the placement was obstructed from the trainee. At this time, I took preliminary steps for the development of the fluoroscopic-based wire navigation simulator under the guidance of Professor Thomas and in collaboration with Professor Anderson. Through iterations of the design, the first preliminary trials were completed in early 2013 using input from Dr. Kho and fellow orthopaedic trauma surgeons Dr. Matthew D. Karam and Dr. J. Lawrence Marsh.

1.5 Significance of the Study

This study details the development and validation of an augmented reality wire navigation simulator. The developed simulator created is an original piece of technology and a novel application of augmented reality. Establishing construct validity of the simulator shows the developed simulator is measuring the skill it was intended to measure. This establishes the augmented reality wire navigation simulator as a viable method for training and assessment of the wire navigation task. Although other orthopaedic drilling simulators exist, this is the first study to show statistical evidence of experienced surgeons outperforming novices in orthopaedic drilling accuracy on a simulator.

This investigation reveals novices often do not account for the inclination angle due to the anteversion of the femoral neck when performing wire navigation. Although this finding is consistent with the conventional wisdom of surgical instructors, this is the first experimental evidence showing statistics on this characteristic behavior. In addition, this work makes preliminary steps towards showing simulator-training transfers to more realistic scenarios. This study documents the learning curve of the wire navigation task on a simulator. These results establish a relationship between wire navigation improvement and number of practice repetitions, a valuable statistic for predicting the practice required to attain proficiency. The learning curve study discovered improvement in the wire navigation task develops more slowly than originally expected, taking multiple days of practice to see noticeable improvement. This study develops a baseline for the learning progression of the wire navigation task that can be a comparative tool for assessing future training and coaching methods.

1.6 Organization of the Study

Chapter Two details the history of surgical simulation and the current apprenticeship model for orthopaedic surgical training. This motivates simulation training through the need for error reduction, the need for more surgical practice, and the potential decrease in cost. To achieve these benefits, a simulator first must demonstrate the simulation improves the skill in the specific task it intends to model. This leads to the validation of an orthopaedic simulator. This chapter also explains the importance of fluoroscopic-based wire navigation as a task heavily influencing the outcome of hip fracture surgery. The development of a wire navigation simulator is motivated by the importance and prevalence of wire navigation in orthopaedic surgery and its potential benefit from increased training opportunities.

Chapter Three executes aim one of the study, describing the design of the orthopaedic wire navigation simulator using augmented reality. Selected stimulus-response cues of the wire navigation task serve as the foundation for simulator development. The documentation includes detailed description of hardware and software components. This section demonstrates the calibration procedures for correlating real world components with the virtual objects. This section also details the operation and additional features of the simulator. The accuracy of the motion-tracking system is assessed ensuring the correlation between virtual reality imagery and real world objects is less than detectable limit for human discernment.

Chapter Four demonstrates the construct validity of the simulator and fulfills aim two of the study. Establishing construct validity shows the developed simulator is assessing the skill it intends to measure, thus advocating simulator training will transfer to a real scenario. The experiment includes forty participants of varying skill levels performing wire navigation on the developed simulator. The experiment measures the wire accuracy, duration, and number of images used in the wire navigation task the simulator can discriminate the skill levels of the participants. The results show experienced surgeons are more accurate than novices, but do not significantly differ in the duration or number of images used to complete the drilling. Novices tend to drill wires inaccurately, biasing towards the posterior region of the femoral head. This probably results from novices not accounting for the inclination angle due the anteversion of the femoral neck. These results establish the simulator as a viable training and assessment tool.

Chapter Five explores the extent of skill transfer from simulator training to a more realistic orthopaedic drilling task. Experimental trials of four trained and four untrained first-year orthopaedic residents were assessed on a more realistic wire navigation scenario using real fluoroscopy. The results showed little to no improvement when trained participants were exposed to simulator training. To further understand this unexpected result, this chapter further explores training novices for the wire navigation task. Through the use of a survey containing radiographic images, the accuracy of first-year orthopaedic residents estimation of the tip-apex distance is quantified. This supplementary investigation shows weaknesses in novices' perception of distances, angles, and anatomical structures in radiographic images. In an additional experiment, a learning curve is constructed using five novices completing the wire navigation task over four successive days of practice on the augmented reality wire navigation simulator, measuring the learning curve for the task.

Chapter Six summarizes the discoveries of this work and generalizes the results for other applications. The proposed aims are briefly discussed. The training effects of the simulator are discussed along with the potential future work concerning the wire navigation task.

CHAPTER 2 – LITERATURE REVIEW

2.1 Introduction

This chapter explores the framework advocating the creation of a wire navigation simulator for orthopaedic training. The history of surgical simulation first highlights how orthopaedics lags behind other surgical disciplines, such as laparoscopy, in the implementation of simulation training. The current training method for orthopaedic surgery is the apprenticeship model, which has been nearly unchanged in the past 100 years. This current method requires trainees to acquire surgical skill through practice on live patients in real operating rooms. Studies of the apprenticeship model also indicate trainees are not achieving surgical proficiency before exiting the training programs. In addition, hospital statistics from the past several decades reveal a shocking number of medical errors resulting from technical complications and lack of skill. This review intends to show the current surgical training method needs reform to decrease medical errors, decrease the financial cost of surgical training, and increase the amount of practice by surgical trainees.

Validated surgical simulators offer an attractive training alternative for reducing cost and risk. A review of orthopaedic surgery shows hip fracture surgery is one of the most common orthopaedic procedures with a high risk for failure due to technical errors. Fluoroscopic-based wire navigation as an important preliminary task heavily influencing the outcome hip fracture repair. A comprehensive literature review of literature reveals few simulators for wire navigation, with even fewer completing validation studies. In addition, all training simulators completing validation studies incorporate haptic-feedback devices, which limit the recreation of the wire navigation task. This literature review points to an unexplored area of developing a validated augmented reality wire navigation simulator without using a haptic feedback device. In addition, the literature review shows no studies examining skill transfer to more realistic drilling scenarios, another area this document will address.

2.2 History of Surgical Simulation

Exploring the history of surgical simulation highlights the underlying need for simulation existence in society. This section explores the realm and history of medical simulators since inception, exploring large variations in price, intent, technology, and effectiveness. The history also reveals the potential for present day surgical simulator development in the orthopaedic surgical field. In addition, the current state of surgical simulation stresses the urgency for orthopaedics to embrace simulation as a possible supplementary teaching method to the apprenticeship model.

Advancements in technology led to the rapid expansion of simulators in the past century. Flight and military simulators dominated the simulator field since modern inception, but current trends show a new diversification of simulators stretching from medicine to business management. Simulation is defined as “a technique, not a technology, to replace or amplify real experiences with guided experiences, often immersive in nature, that evoke or replicate substantial aspects of the real world in a fully interactive fashion” [10].

Simulation emerges from necessity. The situation may entail the need to train a group of pilots without the availability of numerous aircraft. A condition may arise a doctor has never encountered in surgery. An investor may want to test a cutting-edge stock-trading algorithm before investing personal funds. All of these situations give rise for the need of simulation.

Doctors and surgeons’ errors often result in severe negative consequences. By this argument, it only seems natural for medical simulators to emerge due to the high-risk/high-cost nature of surgery and medicine. However, medical simulators lag behind other technical high-risk professions. Ziv et al. hypothesized the complexity of modeling the human body, lack of proof of training effectiveness, and resistance to change in the strong medical culture all contribute to hindrance of medical simulators [11].

While present day medical simulators are less developed compared to extremely high-fidelity immersive flight simulators, the roots of medical simulators extend much farther into history. The *Sushruta Samhita*, an ancient Sanskrit manuscript, was transcribed, revealing to be an ancient textbook of surgery and medicine dating back between the 4th and 6th centuries B.C. The text contained several passages relating to simulation. These simulation procedures created an environment where surgeons became faster and more proficient at surgery. The few anesthetics available at the time made speed essential to successful surgery. Simulation tasks allowed surgeons to increase proficiency and speed. A list of simulation procedures in the *Sushruta Samhita* include incision practice on fruits, suturing practice using cloth, and extracting kernels from jackfruit to simulate pulling teeth [12].

The first trauma simulator emerged from desperation in the summer of 1559. King Henry II of France participated in a jousting tournament and splinters from a shattered lance entered his eye socket. Master Surgeon Ambroise Paré, the most famous surgeon of France, gave little hope for the king. In an attempt to find a solution, Queen Catherine hastily ordered four criminals to be beheaded and required splintered truncheons plunged into their eyes at the appropriate angle to recreate the injury. Unfortunately for the king and queen, this simulation procedure produced no cure, and the king perished 11 days after he sustained the injury [13].

In 1870, the first surgical training simulator emerged, developed by Dr. Benjamin Howard for a new method in treating hernias. The simulator was advertised to illustrate the “descent and protrusion” of hernias at a price of $12 (approximately $210 today) [14]. Dr. Howard emphasized the problem of past and present teaching methods in surgery by stating, “students are more often qualified to answer questions upon it, than to treat it.”

Howard further mentioned the difficulty of obtaining and keeping cadavers for teaching purposes [15].

The first widely successful medical simulator emerged in the early 1960s as Resusci-Anne, a manikin for cardiopulmonary resuscitation [16]. This simulator was a consequence of Dr. Peter Safar's finding of the effectiveness of mouth-to-mouth and mouth-to-airway methods for artificial respiration [17]. Asmund Laerdal, a Norwegian toymaker, made the first iteration of Resusci-Anne after encouragement from Norwegian anesthesiologist, Dr. Bjorn Lind. The model encouraged hyperextension of the neck to open the airway before insufflating the lungs. Later versions of Resusci-Anne contained a spring inside the chest for compression routines [18]. CPR training courses still presently use the mannequin.

Not long after Resusi-Anne, researchers and physicians created more advanced high fidelity simulators than previously existed. The goal of the era was to recreate all the human vitals in a simulator with different diagnoses conditions. The most successful of the era was "Harvey," a mannequin device used to mimic a cardiologist patient. The 1976 prototype was capable of "reproducing blood pressure, jugular venous pulsations, carotid and peripheral arterial pulsations, peripheral arterial pulsations, precordial impulses and auscultatory events of almost all cardiac diseases" [19]. Harvey is still presently being used to teach novice residents methods of taking blood pressure and recognizing heart murmurs, and senior residents the ability to diagnose 27 unique cardiac conditions. Many of the versatile high fidelity simulators range from $75,000 for Harvey to well over $200,000 for advanced anesthesia simulators [20].

One of the most successful modern day simulators is the Minimally Invasive Surgery Trainer – Virtual Reality (MIST-VR). This device entered the medical field in 1997 as both a training and assessment tool for laparoscopic surgery for supplementing the apprenticeship model. The trainer contained two laparoscopic instruments and a virtual abstract shape for pick-and-place and manipulation [21]. The landmark validation study of surgical simulators, demonstrating simulated training improves performance, was published by Seymour et al. in 2002, showing the MIST-VR simulator training decreased the number of errors in laparoscopic cholecystectomy [22]. Extensive successful validation studies of this device drove to the device to be widely adapted leading to its ultimate success [23-26].

With the success of the MIST-VR the field of laparoscopic surgery simulation increased. The cause of the early inception of laparoscopic trainers attributes to the instrumentation used during laparoscopic surgery. In general, standard laparoscopic handles fit with a variety of modular tool tips for surgery completion. Unlike other surgeries, the input handles are standard resulting in easy implementation of a variety of tools into a virtual system. This inherit feature of laparoscopy allows nearly complete procedures to be simulated in a virtual environment without the need to track and develop hardware for every single tool. In addition, the long handle laparoscopic tool dampens the haptic information passed to the surgeon. Since haptic feedback is not as prevalent as in other surgeries due to the long instruments, it is often neglected in laparoscopic simulators. This further simplifies the development process of a laparoscopic simulator.

Contemporaneously to MIST-VR, the MISTELS (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills) simulator emerged using a lower cost, lower-fidelity approach to training methods. The simulator uses actual laparoscopic tools inside a box to obstruct the user’s direct view of the workspace. A camera attached to the inside of the box output an image to a video monitor giving the participant visual position of the tools. Pattern cutting, knot tying, mesh placement over a defect, placement of a ligating loop, clip application, and pegboard patterns make up the seven individual exercises for practice [27]. Similar to MIST-VR, MISTELS underwent validation studies [28, 29].

The success of the MISTELS system acquired recognition from the Society of American Gastrointestinal Endoscopic Surgery (SAGES) and formed a new committee in charge of developing educational materials for fundamentals of laparoscopic surgery (FLS). The curriculum consisted of two components: cognitive and psychomotor skills. For the psychomotor skills, FLS adopted a modified MISTELS system with only five of the original seven practice exercises [30]. The adaptation and requirement of simulation has excelled laparoscopic simulation to the forefront of medical simulation.

Many surgical disciplines are adding simulation as a required facet to their curriculum. Advanced medical education residences at institutions such as Emory, Harvard, Johns Hopkins, Mayo Clinic, Northwestern, and Stanford University are supporting fully-equipped medical simulation centers [31]. However, the question remains as to whether orthopaedics will be at the forefront of the paradigm shift or one the final programs to embrace change into the simulation generation. Simulation will likely never replace current surgical teaching methods, but it is important to utilize simulation techniques proven to increase proficiency. The developments of technology promote the continued augmentation of the current teaching methods with validated alternative methods.

2.3 Teaching Techniques of Surgery

Teaching is the art of transferring knowledge or skill from one person or source to another. Due to the individual nature of learning, no single “superior” method for teaching exists. Often, several teaching methods must be employed to target all types of learners. A majority of teaching in university settings is lecture-based combined with assignments and projects promoting both independent and team learning. Among trade professions, a higher majority of learning occurs within hands-on and apprenticeship models. Selection of teaching method may be intensely challenging when the subject encompasses a significant range and variety of cognitive and skill-based components, such as those included in surgical training. In fact, surgery has been described as “…the most complex psychomotor activity that human beings are called to perform” [32].

Surgery performance combines medical knowledge, planning, judgment, dexterity, communication, and teamwork. Such a large domain necessitates education development that is practically as versatile as the occupation itself.

To design and meaningfully improve training methods in surgery, one must comprehensively explore the history of surgical training as well as the current state of surgical education. Currently, there are no accepted assessment procedures for screening doctors into surgical or non-surgical fields. The strengths and weaknesses of current training methods can be examined through statistical analysis focused on the underlying essential abilities of surgeons. By dissecting surgical abilities to intelligence, visual-spatial skill, and motor abilities, one can cross-examine literature to see the impact of each aspect on surgical skill. Finally, through correlating cognitive models of learning to surgical abilities, along with experience and practice, one can make a final assessment on further advantages to future surgical education paths.

2.3.1 Brief History of Surgical Training

To understand the current state of surgical education, one must first explore the history and development of modern surgery. Surgical training was a historically direct product of each surgeon's unique and varied experience and expertise. Two categories of surgeons existed: those academically trained, and so-called "barber-surgeons," who typically possessed little or no training. However, the general public was virtually incapable of deciphering well-trained surgeons from inexperienced surgeons [33].

Until the latter part of the 19th century, a majority of surgeons remained self-taught, leaving professionally qualified surgeons few and far between. Regardless of setting, aspiring surgeons obtained surgical knowledge and training nearly exclusively from sheer experience and collective encounters. The few professionally qualified surgeons in existence were not interested in teaching, but more focused on developing their own private practices and skill. William Stewart Halsted, an American surgeon, critically observed the existing deficiencies of the surgical teaching model and devised a new program at Johns Hopkins Hospital. Halsted had learned through observation and collaboration with experienced surgeons over a two-year period spent in Europe. He postulated that a similar program could greatly improve the deficient state of surgical care in America. Subsequently, he structured a similar, formal educational model at Johns Hopkins, showcasing a skill development model facilitated by continuous practice, feedback, and critically timed instruction. Halsted's training model was widely revered throughout the medical community and quickly adopted throughout hospitals in Baltimore. In short order, the model expanded throughout the United States [34].

The original training period for a resident was 8 years, including 6 years as an assistant and 2 years as the house surgeon. Halsted had 17 chief residents throughout his tenure at Johns Hopkins. Eleven of the 17 chief residents moved on to form their own residency programs throughout the country, a change from the prior trend of entering private practice. This method preserved prior knowledge while new knowledge was developed and taught to future generations [34]. "See one, do one, teach one" became the common phrase used to explain the apprenticeship model. The apprentice-based teaching model, sometimes referred to as the Halstedian teaching model, still serves as the primary method for surgical education.

The longevity of the apprenticeship model demonstrates how the medical community has had overall success in the sophisticated development of both basic and technical surgical skill. Thoughtful dissection and analysis of the acquisition skill models reveals both why the apprenticeship model is successful and areas where improvement is possible. For example, the years of surgical training develop technical skill by placing residents in real scenarios, allowing abundant opportunities for practice, with instructors providing timely feedback. These components are the strengths of the apprenticeship model.

2.3.2 Effects of Contributory Characteristics in Surgical Skill

Talented individuals are often referred to as “gifted” or “a natural” with respect to their skills because talented people appear to have been born with innate, specific traits enabling easy success in a specific field. Surgeons are commonly accredited with unusually “steady hands,” implying perhaps an innate ability, and indicating superior skill over an average human being. While some characteristics such as height and weight are hereditary and tend to be valuable in certain athletic endeavors, an age-old debate has attempted to determine whether hereditary traits expand to technical skills and occupations.

Disciplined analysis of skill acquisition, relative to a particular occupation, is imperative to moving toward the identification of certain prerequisites to success in a given occupation. In medical education, identifiable metrics can be used to separate medical students who have the talent to become surgeons from others whose measurable skills indicate a lower likelihood of success in a surgical program. However, before performing an analysis, the specific contributory traits must be correctly identified as a general ability that contributes to surgical skill. Studies typically explore statistical correlation between skill and general aptitude testing. After a correlation study, surgeons may then be evaluated utilizing a variety of criteria, including intelligence, visual-spatial ability, and psychomotor ability compared to less-skilled individuals.

Despite detailed study, the question still remains: are proficient surgeons naturally gifted with dexterity and spatial skills or perception, or is surgery an acquired skill, built through repetition and practice? Early accounts and speculation attributed greatness to divine gifts [35]. Even as science progressed in the 19th century, the predominating view was that expertise was inherited or innate. Specifically, Sir Francis Galton attributed talent to natural “ability, zeal, and the capacity for hard labor” [36]. Modern views in behavioral genetics describe expertise as a result of “interactions between environmental factors and genes during the extended period of development" [35]. However, it is not clear how much expert performance is dependent on genetic composition, and how much is dependent on environmental factors, such as practice. Unfortunately for realm of surgery, most of the body of evidence correlating heritable characteristics with elite performance is in sports, where anatomical characteristics, such as height in basketball players, differ greatly from the average population. Although these anatomical characteristics might lend themselves to an advantage, there is insufficient evidence support the theory innate ability or skill is created by hereditary factors.

Edwin E. Ghiselli completed extensive research in the underlying abilities of various occupations for predicting occupational success. He examined two different types of occupational criteria: training and proficiency. Training included the capacity to acquire knowledge and skills essential to the performance of the occupation. Proficiency pertained to the level of achievement in the occupation practiced by workers who had already been trained. Ghiselli found a strong positive Pearson's coefficient correlation between intellectual (0.41) and spatial/mechanical abilities (0.41) for predicting trainability in trade and craft occupations. A moderate positive correlation (0.31) for perceptual accuracy and a weak positive correlation (0.17) for motor abilities was found in predicting the ability to train novices. An insignificant negative correlation was found in the trainability based on personality (-0.13) [37, 38].

Correlations for predicting competence in the trade and craft occupations were much lower than predictions of trainability of subjects for most categories. Intellectual abilities, spatial/mechanical abilities, perceptual accuracy, motor abilities, and personality traits had Pearson's coefficients of 0.19, 0.23, 0.22, 0.22, 0.19, and 0.25, respectively. This was the general trend across most occupations, with predictors for trainability higher than competence predictions [37]. This trend points to the existence of underlying abilities and intelligence to acquire the necessary skills and knowledge to perform an occupation, but shows it does not predict the competence of the human at the occupation.

One can infer that specific occupational skill cannot be detected through general tests and exercises, and that it is specific to the occupation. General aptitude tests determining admission into an occupation or residency program seem to be a more plausible use than detecting later competence in an occupation. For instance, higher visual-spatial scores on aptitude tests relate to higher initial competency in a surgical task, but the effect is neutralized after training and practice [39]. This type of scenario leads to the strong argument that specific acquisition of desired skills in an occupation is most likely due to repetition and practice of specific occupational procedures.

Specific to surgical fields, Lee et al. studied and compared the motor dexterity of medical students pursuing surgical fields to students pursuing non-surgical fields. Results found no significant difference between the two groups when measured using the Purdue Pegboard test [40]. This suggests doctors pursuing a surgical career do not have an innate motor ability superior to the average medical student.

Wanzel et al. studied visual-spatial, manual dexterity, and surgical performance among dental students, surgical residents, and staff surgeons. There was no statistically significant difference in visual-spatial ability or manual dexterity between the three groups, but there was a highly significant difference in performance in a bench model fixation of an anterior mandible. Surgeons outperformed residents and residents outperformed the dental students [41]. When comparing trainees in surgery, psychiatry, anaesthetics, and medicine results, the results showed no significant difference in psychomotor skills or visual-spatial ability between the groups [42]. To further complicate the matter, drawing a conclusion on underlying abilities, one study found master surgeons had lower visual-spatial ability than medical students [43]. Physicians medical school grades 17 years prior also show a weak correlation (around 0.2) for occupational success [44].

Extensive research shows that hereditary factors are very limited in their contribution to surgical skill. Also, simple cognitive or manual tasks are unable to predict surgical skill. The inability to predict performance of skill using cognitive and manual tasks suggests that skill learning is task specific in the medical field, suggesting that the practice of simple dexterity tasks will improve surgical skill very little. This points back to the strengths of the apprenticeship model. Surgical residents learn specific applied skills used in surgery in a real scenario. These specific environmental conditions are maturing and developing residents into proficient surgeons.

If skills are not based on hereditary, intelligence, or innate abilities, what are the specific environmental factors causing the development of skill? The most highly regarded environmental factor reflecting exceptional performance is practice. K. Anders Ericsson states, “The central mechanisms mediating the superior performance of experts are acquired; therefore acquisition of relevant knowledge and skills may be the major limiting factor in attaining expert performance” [45]. The effects of fundamental elements and practice influence expert skill must be examined to further understand the best methods for training surgical residents.

2.3.3 Deliberate Practice in Surgical Training

The inclusion of increased practice is one component of the apprenticeship model attributed to its past success. To explore the impact of increased practice, an exploration and dissection of medical studies must show that increased experience, practice, and volume correlate with better medical outcomes. Skill acquisition and expert performance across various disciplines are reviewed supporting deliberate practice as the key ingredient to attaining expertise. Lastly, a deliberate practice model specific to medical education is summarized.

Under the apprenticeship model, residents obtain hands-on skills by first watching the technique then performing the technique. Analysis of patient statistics shows that increased practice is beneficial to clinical outcomes. Luft et al. started volume-outcome research in this area in the late 1970s into the 1980s, showing a strong negative correlation between the volume of specific procedures performed at a hospital and mortality [46-48]. Further studies correlate high volume with better outcomes for surgical patients, but mixed evidence for general medical patients [49, 50]. In addition, a strong correlation existed between higher volume and better outcomes for relatively complex surgeries, usually performed on an elective basis [51]. This data suggests that an increased volume of trials is more important for surgeons performing complex surgeries than medical clinicians performing routine procedures. Furthermore, the data suggests that the more complex a procedure, the more practice is required to become proficient.

In a detailed study of radical prostatectomy, Vickers et al. observed a steep surgical learning curve, without reaching a plateau until 250 performed operations [52]. Another study showed less post-operative complications due to a radical prostatectomy completed by a high-volume surgeon at a high-volume hospital [53]. For a mildly complex surgery, this data demonstrates that practice acquired through volume of procedures completed may play a major role in the outcome of surgery.

While volume plays a strong role in complex surgeries, it also impacts other common medical procedures. In a study of hip fracture surgery, Bjorgul et al. found that residents require 20-30 trials before reaching the speed of an expert during a procedure. This study also noted that learning curves vary widely for different residents [54]. This data suggests large amounts of practice and many years may be required to reach mastery of a single procedure alone. However, it also shows that each resident is unique in terms of time required to become proficient. In addition, it is difficult to assess which developed skills will transfer across multiple procedures.

Similarly, Taylor et al. found a significant negative relationship between volume and mortality for major orthopaedic surgery, including total hip arthroplasty, partial hip arthroplasty, revision total hip arthroplasty, total knee arthroplasty, and revision total knee arthroplasty [55]. To further advocate the need for practice, Bianco et al. studied effects of surgical fellowships. The results indicate that surgical fellowships enhance overall surgical outcomes and speeds up the rate of learning [56]. This lends additional merit to the idea that surgeons need to perform a large number of surgical procedures to obtain proficiency.

While these strong correlations support the notion that larger volumes of treatments result in better surgical and medical outcomes, there is little evidence that explains the precise causal link between practice and outcomes. One hypothesis states that outcomes are better because at larger volumes, more practice occurs, resulting in increased proficiency. This hypothesis is often called the “practice-makes-perfect” model. The other hypothesis states because the clinicians are better, they encounter more referrals, known as “selective-referral patterns,” resulting in better outcomes. Halm et al. completed the most comprehensive methodological study of the volume-outcome relationship in 2002. The study reviewed 135 volume-outcome studies spanning over 27 procedures and clinical conditions. The authors confidently asserted a robust correlation between higher volume and better health outcomes. However, differences in outcomes between high and low volume treatment centers were much more modest for common procedures where selective referral and regionalization policies have been proposed. These procedures include coronary artery bypass grafting (CABG), coronary angioplasty, and carotid endarterectomy [6]. These empirical results suggest that increased practice results in better medical outcomes.

A plateau or an asymptotic behavior is often observed after continual repetition, tending to level off toward what is believed to be a maximal performance level. This point is often thought to be the pinnacle of what an individual can attain, yet it is little more than a common misconception. Plateaus in the learning curve are likely a cause of other factors, and experiments have shown that a plateau period during skill acquisition can be mitigated such that further improvement can be achieved through better and different training methods [57]. For example, in 1896 when athletic training was still in its adolescence, the Olympic games marathon winner, Spyridon Louis, completed the marathon in 2 hours 58 minutes. In 2013, a man above the age of 60 completed Boston Marathon in a faster time [58]. Such a difference in results can be attributed to better training methods in present day athletics.

In the analysis of proficiency, the first question is typically, “How much practice is required to obtain expertise?” In 1973, Simon and Chase first recognized that it took ten years to develop a mastery of chess [59]. Bobby Fischer and Salo Flohr both required nine years to obtain master status in chess [60]. Scientists on average produce their greatest work at 35.4 years of age, whereas poets and authors produce their greatest works at the age of 34.3. These ages are correlated with approximately ten years after their first publications. First publications for scientists average at the age 25.2 years old, compared to 24.2 years of age for authors. This ten-year rule has also been applied to other domains such as music, mathematics, and swimming [35].

The existence of the ten-year rule raises another question: if it takes only ten years of practice to become an expert, why are we not experts at our everyday tasks? By the ten-year rule, people driving in and out of the city five days per week for work should be as good as competitive racecar drivers. People with desk jobs should become master typists. The average jogger should progress to a world-class runner.

The reason we are not masters of daily tasks is because as our skills increase to a satisfactory level, the task becomes autonomous to a degree, requiring very little cognitive input to complete the task. Once the requisite cognitive effort ceases, improvement also slows. This explains for example why people write less legibly than they are capable, answer simple questions more slowly than mentally achievable, and generally perform everyday tasks at a casual pace [61].

Once the autonomous stage is achieved, the learner must develop techniques to continually present a challenge. Increasingly challenging mental representations of a task allow performers to continually improve over time. Some performers may reach an advanced level and grow comfortable with their level of achievement and therefore stop continuing to improve. This stage is termed "arrested development." This is seen in musicians who reach a stage where performing is effortless in their realm and further improvement is not necessarily desired to achieve greatness. A graphical representation of this phenomenon is shown in Figure 1.

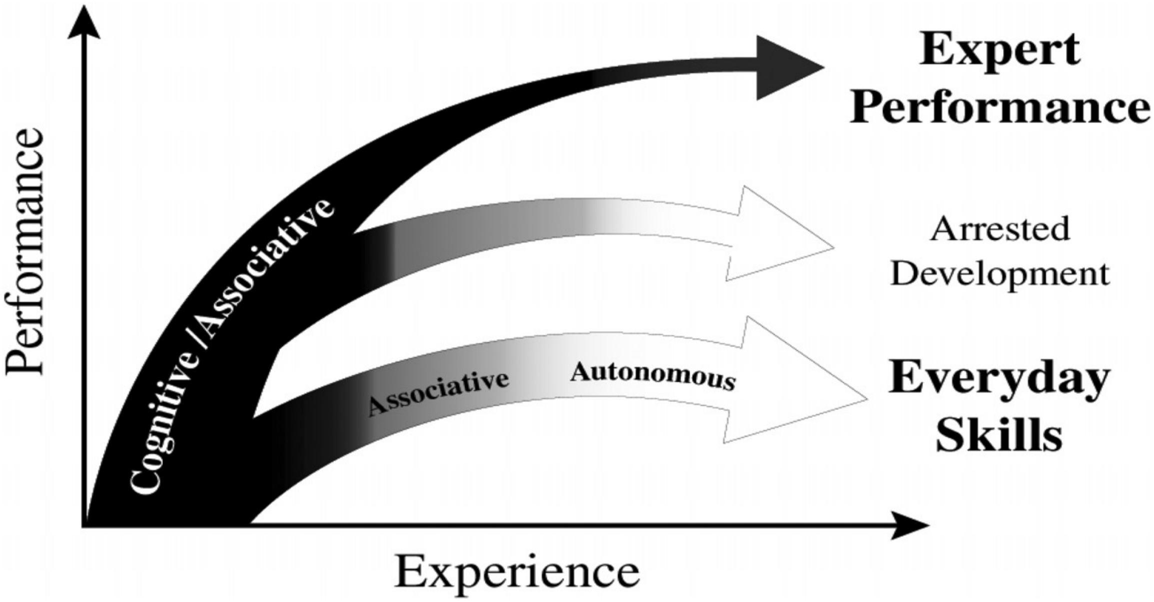

Figure 1. Showing cognitive effort and experience as a relationship to cognitive effort and performance. Demonstrating arrested development. Graphic from [62].

While a surgeon spends well over ten years in training, is that enough time? After completing an undergraduate degree, the future surgeon is in medical school for approximately four years. After medical school, the future surgeon completes a postgraduate residency of four or five years. By the time the surgeon has finished training, he or she has completed approximately twelve or thirteen years of training. However, which specific activities contribute to the ten years required to attain expert skill? Some may argue that undergraduate study does not enhance surgical ability, meaning that new surgeons have fewer than ten years of experience, in turn meaning that they are not experts in their field. Delineation of the specific activities contributing towards the attainment of mastery is most certainly needed, along with a better refinement of the ten-year rule.

Performance improvement does not merely depend on increased repetition to further skill development. Practice should not be confused with increased repetition. The key to consistent improvement is “deliberate” practice. Expert skill does not develop from solely a specific time frame of repetition; there must be cognitive effort for constant improvement. This cognitive effort is arguably the most important aspect to improving. In addition, subjects must receive immediate, informative feedback while trying to improve, with the particular type of feedback matching the characteristics of the task. Finally, subjects should perform the same or a very similar task repetitively to increase performance [63].

Ericsson, Krampe, and Tesch-Römer proposed a theoretical framework for the acquisition of expert performance through deliberate practice. Through review, Ericsson et al. assumed and hypothesized the amount of time spent in deliberate activities is monotonically related to individual performance. On this assumption, a framework was established to develop expertise. First, deliberate practice requires time and energy, accompanied by access to teaching and training material, including expert instruction. Second, an individual’s engagement in deliberate practice must be motivating, at least to a degree where the individual continuously yearns to improve. This exists as intrinsic motivation or inherent reward from the enjoyment of the result of deliberate practice. The performer does not necessarily need to enjoy the act of deliberate practice to enjoy the result of increased performance. This suggests that motivation may culminate from either the enjoyment of the practice, such as a person who loves playing music, or enjoyment may come from the end result, such as a person who enjoys winning many awards from continually practicing music, but who does not necessarily like to practice. Finally, deliberate practice must be an effortful activity only maintained for a limited period each day. Exceeding this limit leads to exhaustion and wasted effort [35].

By using the framework of deliberate practice, Ericsson et al. further examined the ten-year rule for obtaining expert performance. It was found that amount of time to become an expert directly depends on the number of hours of deliberate practice, not necessarily a set timeframe. By estimating the number of deliberate practice hours by violinists at the age of twenty, a standard was set for expert performance at 10,000 hours of deliberate practice. At the age of twenty, the most advanced group of expert violinists accumulated 10,000 hours, while the next most accomplished group had 2,500 fewer hours. The least accomplished group of experts accumulated 5,000 hours of deliberate practice [35]. This is shown in Figure 2.

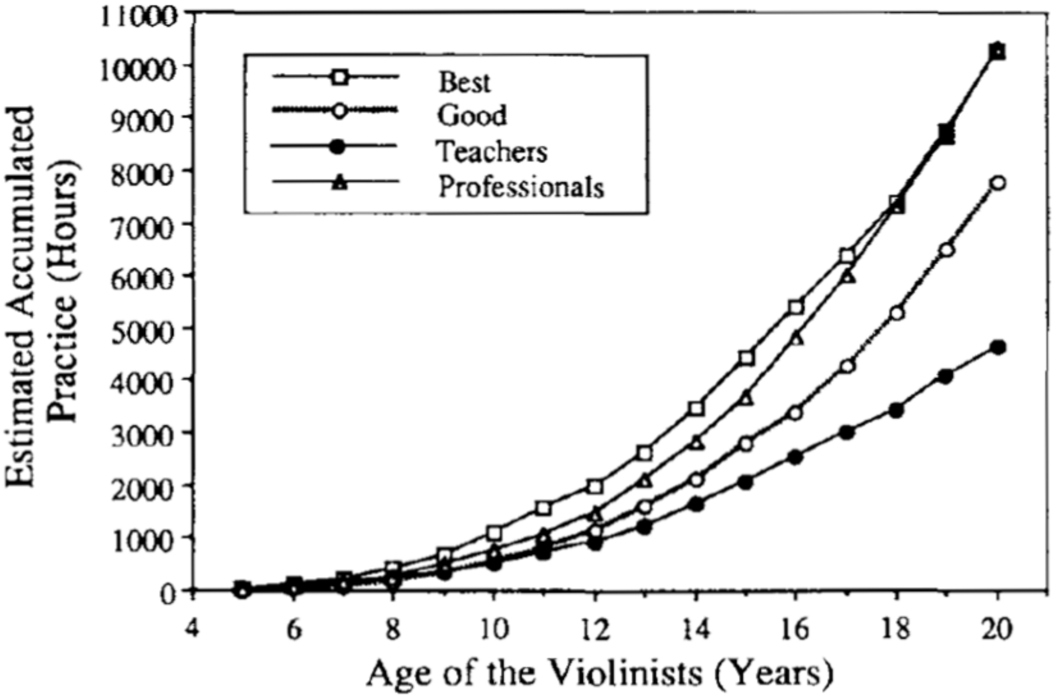

Figure 2. Relation of the level of expertise in violinists and accumulated hours of practice, from [35].

The domains of music, sports, and chess are well explored and proven for the methods purposed by deliberate practice. To fully prove the strength of deliberate practice, it must be explored in the medical domain. Butterworth and Reppert examined medical students, residents, general practitioners, and cardiology specialists in their ability to identify heart conditions by recorded sounds. The results showed a bifurcation in performance based on experience. The accuracy of the diagnosis increased with training for medical students, residents and certified cardiologists. General practitioners showed a decrease in accuracy with increased experience, showing increased degradation since the end of medical training [64]. Increased performance ceases with general practitioners not being required to diagnose heart sounds as often as other medical students, residents, or cardiology specialists. This is shown in Figure 3.

In a training situation of a surgical drilling task, residents who distributed practice over four weeks outperformed the residents who attained all their practice in one day [65]. The benefit of spaced practice over massed practice is shown over many domains [66]. This advocates for training throughout a curriculum, not merely learning a single skill in an all day skills session. A study performed by Pusic et al. on deliberate practice asked residents to identify an abnormality in ankle fracture radiographs. When shown a bank of individual radiographs, the resident was to declare if there was an abnormality and where it existed. After the resident determined their answer, immediate feedback was given. This study resulted in an 20 trials before the subject reached an efficient phase of learning, and an addition 30 trials before learning started to plateau [67]. Degradation of skill shows deliberate practice is needed not only for improvement but also for skill maintenance after formal medical training.

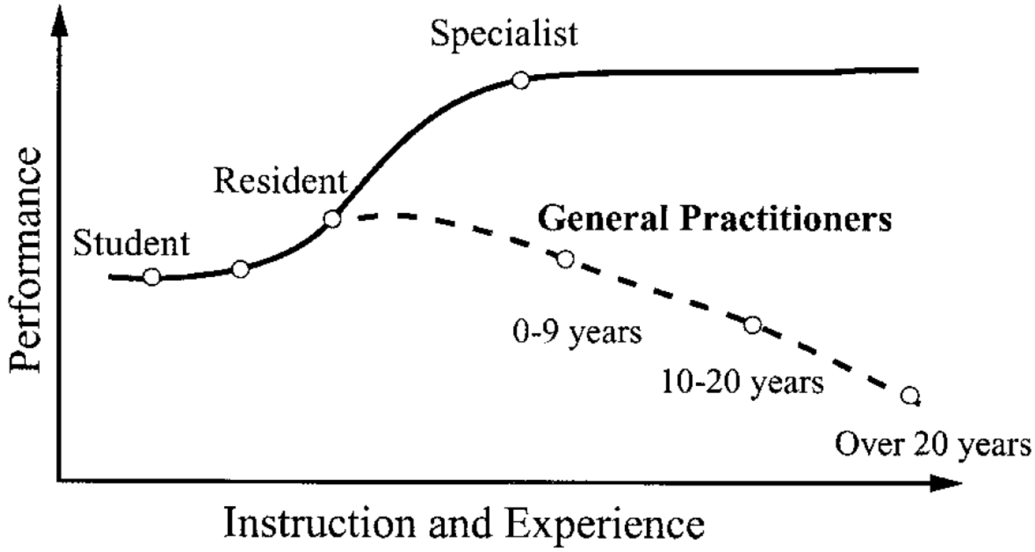

Figure 3. Bifurcation of performance with increased experience, from [68].

An analysis of deliberate practice methods combined with simulation-based medical education methods was completed. Of the 14 studies reviewed, results show the deliberate practice using simulation-based techniques proved to be an overall superior method of training over the apprenticeship model in the medical field [69].

McGaghie et al. adapted Ericsson's framework and established a list of nine requirements for continuing medical education for undergraduate and graduate medical students. These ideas are to encourage constant improvement of skills rather than merely maintaining acquired skills. The skills are shown in Table 1.

Table 1. Deliberate practice requirements defined by McGaghie et al. for continued medical education [70].

1. Highly motivated learners with good concentration
2. Engagement with a well-defined learning objective or task
3. Appropriate level of difficulty
4. Focused, repetitive practice
5. Rigorous, precise measurements
6. Informative feedback from educational sources (e.g., simulators or teachers)
7. Monitoring, correction of errors, and more deliberate practice
8. Evaluation to reach a mastery standard
9. Advancement to another task or unit

Deliberate practice tends to be the one commonality across all domains that improves performance. The emphasis on the training remains that the practice must be "deliberate." By merely repetition alone, a plateau will occur as soon as the task does not require additional cognitive effort and becomes autonomous. To develop better training methods in training surgical skill deliberate practice is necessary for advancement. To further understand why deliberate practice is essential to skill development the following will explore some of the most widely recognized learning models for skill acquisition.

2.4 The Need for Simulation

Despite the cemented role of the apprenticeship model in surgical training, the model has many pitfalls and disadvantages. Building on the previously explained aspects of skill acquisition, including deliberate practice and attainment of expertise, this section makes a case for supplementary training in the surgical realm via simulation. This section will also explore the current state and fidelity of simulation technology. In addition, the few validated simulators will be analyzed to provide an educational backdrop for creating and implementing new simulators into surgical curricula.

2.4.1 The Need for Error Reduction

> Ladies and gentlemen, welcome aboard Sterling Airline's Flight Number 743, bound for Edinburgh. This is your captain speaking. Our flight time will be two hours, and I am pleased to report both that you have a 97 percent chance of reaching your destination without being significantly injured during the flight and that our chances of making a serious error during the flight, whether you are injured or not, is only 6.7 percent. Please fasten your seatbelts, and enjoy the flight. The weather in Edinburgh is sunny [71].

The quote above applies the risks of medical errors to the airline industry. If the risks of medical procedures were applied to flying, the public would be unwilling to accept them. Why does the public accept the current risks in the medical field?

Errors are responsible for a tremendous degree of patient injuries and deaths. They are events readily understandable to the public, and whether they result in injury or expose the patient to the risk of injury, are events that are clearly desirable to prevent. A significant body of risk-mitigation knowledge, paired with highly successful experiences in other industries provides an excellent backdrop for addressing the safety problems of the health care industry. Finally, while the rapid development and evolution of health care has most certainly led to improvements, it also holds the potential to create new hazards.

Understanding the cause of medical errors requires investigation into specific research studies. McGuire et al. completed a 13-year study, statistically analyzing 44,603 major operations. Of all the surgical complications experienced in these operation, over half resulted from documented medical error. Approximately eighty percent of all medical errors occurred during operations [72]. Thomas et al. estimated that 64,809 Americans die each year from medical management errors [73]. Yet another study, based in New York, estimated that 98,609 Americans die due to medical errors every year [3]. This statistic ranks medical errors as the 6th leading cause of death in the United States, as shown in the mortality statistics from the 2010 CDC report in Table 2 [4].

Table 2. Projection of medical error mortality [3] into top causes of mortality in the United States [4].

Cause of Death	Number of Deaths
1. Heart Disease	597,689
2. Cancer	574,743
3. Chronic lower respiratory diseases	138,080
4. Stroke (cerebrovascular diseases)	129,476
5. Accidents (unintentional injuries)	120,859
6. Medical Errors	**98,609**
7. Alzheimer’s disease	83,494
8. Diabetes	69,071
9. Nephritis, nephrotic syndrome, and nephrosis	50,476
10. Influenza and Pneumonia	50,097

Technical complications were the third-highest cause of adverse events (13 percent), just behind drug complications (19 percent) and wound infections (14 percent). Few technical errors were attributed to negligence compared to other adverse event categories [5]. The distinction between omission and negligence is noteworthy. Negligence exists when a degree of error exceeds a given accepted norm. Omission is the failure or delay to institute necessary medical action in a given scenario. An argument can be made that because a lower percentage of technical errors occur due to negligence, more errors are made due to lack of technical skill. In fact, an additional study of 45 surgeons at three teaching hospitals revealed that two-thirds of all errors occurred during the intraoperative stage of surgical care [74]. This statistic also suggests that lack of critical skill in the operating room is the source of medical errors.

The Institute of Medicine identifies medical errors as an immediate problem. In their eye-opening report on errors in the medical industry titled, “To Err is Human,” they proposed five principles for the design of safety systems in health care organizations. The fifth principle details “creating a learning environment for safe health care organizations.” At the core of this principle is “Use simulation whenever possible” [75]. Simulation allows for the development of novice practitioners through a safe skill development model when new or hazardous procedures are introduced.

2.4.2 The Need for Surgical Practice

The assistant executive director of the American Board of Surgery, Richard H. Bell, Jr., estimated that the number of hours an average resident spends in deliberate practice is 1,148 hours over a five-year period, considerably short of the requisite 10,000 hours needed to attain expert level. In his address, the lack of technical skill by surgeons, or as he terms “Why Johnny cannot operate,” highlights the need to supplement current training programs with valid simulation exercises to better prepare surgeons. Bell, Jr. goes as far as recommending moving “the simulation agenda forward with a national consortium.” This entails approaching the U.S. Congress and asking for a substantial investment [32].

In a detailed study, program directors concluded that 121 out of 300 general surgery procedures were deemed essential to the practice of general surgery. Examining 1,022 graduating residents, these residents performed, on average, only 18 of the 121 procedures more than ten times throughout their residency. Thirty-one of the 121 procedures were performed less than once on average [7]. Raphael S. Chung commanded a study of residency programs in the early 2000s. This study showed that an average resident spends approximately 3,963 total hours in the operating room throughout a five-year residency. While that estimate is still far short of 10,000 hours, it is further disheartening that the estimate encompassed the mere total number of hours in an operating room, not necessarily hours of a deliberate practice. A resident may only be observing during this time and not individually performing any surgical task.

In addition to practice purposes, surgical programs are searching for new, more efficient surgical training methods due to the mandatory reduction in residents' work hours. In 2003, the Accreditation Council for Graduate Medical Education (ACGME) mandated a reduction of duty hours to a maximum of 80-hours per week. Since mandatory reductions were implemented, studies have shown mixed results when attempting to quantify the impact of reduction of hours by residents [76-78]. The mixed results could be attributed to how strictly or uniformly schools abide by the 80-hour mandated rule, since a certain degree of noncompliance has been observed [79]. The study results also helped to push some residency programs to create innovative methods to reach compliance [80].

Before placing blame on the ACGME for cutting work hours, the routine training day of a resident must be examined. The 80-hour workweek allows for nearly 11.5 hours of work per day, working seven days per week. Studies show expert performers typically never exceed four to five hours of deliberate practice per day. Exceeding this level can lead to "burnout" and reduce improvement effects. This holds true across a variety of domains from music performers to athletes and writers. The limiting factor is the required concentration to make practice deliberate, as it is difficult to sustain concentration for more than five hours. However, experts have little trouble sustaining four hours of practice per day for numerous years [35, 81]. Applying the current estimates of deliberate practice to surgical training, residents have been shown to spend between a half hour and an hour and a half of deliberate practice per day [32, 82]. These deliberate practice times are nowhere near the amount of practice per day to acquire expert skill, yet an 80-hour workweek well exceeds the threshold for exhaustion.

Therefore, the problem does not lie within the reduction of hours to 80 hours per week. In fact, the problem is lack of efficient training during each 80-hour time window. For example, a major league baseball player comes up to bat on average three to four times a game. During each at bat, the player sees an average of four pitches. This means in a game that averages nearly three hours, a batter sees roughly 12-16 pitches. If a batter is looking to improve his hitting, playing game-after-game is a very inefficient way to practice. There are simply not enough hours in a day to reach the expert level. Instead, coaches place hitters inside a batting cage, where they can hit 50 baseballs in five minutes, the equivalent of playing four consecutive games. Dissecting the task into several important factors allows the subject to attain reasonable practice in a short period of time.

This logic may also be applied to surgery. Surgery takes a few hours, including preoperative procedures and the actual surgery. A surgeon may only get to complete suturing once during an entire surgery. If a surgeon is trying to reach an expert level of suturing, it is an inefficient way to attain practice. This is one major downfall of the apprenticeship model, as the practice is time-consuming and inefficient. It is inefficient to “play the entire game” to attain skill in a specific area. The apprenticeship model is very good at exposing students to a variety of scenarios and situations, but it lacks in the areas for attaining technical skill. Simulation could greatly supplement the apprenticeship model by enabling students to gain a vast amount of technical practice in a short period of time. Instead of acquiring a mere hour of deliberate practice a day, a resident could easily double the amount of deliberate practice per day using simulation for practice.

2.4.3 Cost of Resident Training

The apprenticeship model is the epitome of “hands-on real world” training, but it comes at a cost. Any increased time in the operating room is expensive, a cost which is passed to the patient. The cost of instruction in the operating room is particularly high because the entire surgical team, ranging from imaging technicians, nurses, to anesthetists, are compounding the expense. The following highlights the individual costs and adds to the need for more simulation in the learning environment.

It is difficult to acquire a simple quantitative measurement of the cost of training a resident in the operating room. Bridges and Diamond quantified the cost by comparing operating times in procedures where residents were present compared to procedures with no residents in attendance. In this four-year study, the average extra time was 11,184 minutes per resident. The cost per minute of extra time was estimated at $4.29 per minute. This estimate excluded supplies, indirect costs, anesthesia costs, and surgeon fees. In total, the results showed that the increased cost of training residents in the operating room was $49,979 over four years [83]. Extrapolating this conservative estimate to a five-year residency, the cost in lost time alone can be estimated to $62,474. Other studies estimate operating time costs much higher, around $66 per minute [84]. Using this estimate, the cost of training a single resident exceeds $738,000.

Similar studies observed the same trend. The increased time during resident training is accentuated by the complexity of the surgery. When measuring the time from skin incision to skin closure (“skin-to-skin” time), a hernia surgery was found to take an increased eight minutes of time, where carotid endarterectomy increased time by 44 minutes and partial colectomies took an average of 60 minutes longer with a resident [85]. An orthopaedic study of ACL reconstruction noted an average increased cost of $661.85 when a resident completed the surgery, compared to a faculty surgeon [86].

Analysis of the cost of resident training further justifies the expense of simulation training. Simulation training before entering the operating room has shown to significantly decrease operation time in laparoscopic surgery. Seymour et al. found virtual reality trained surgeons were faster and less likely to make an error on a laparoscopic gallbladder surgery [22]. Torkington et al. revealed subjects trained on laparoscopic simulators had faster left hand movement and fewer overall movements when subjected to a different laparoscopic assessment device [25]. Fried et al. observed superior cutting, clipping, mesh placement, and suturing with a single group trained on a simulator and tested on a live animal model [29]. Andreatta et al. observed increased speed and accuracy on a porcine model simulation after training on a virtual reality laparoscopic simulator [87]. Further, the decrease in operation time and errors largely effects total operational cost. Often times, simulation systems are dismissed due to their high initial cost. However, closer examination of long-term savings often concludes that it is a feasible option for investment. Depending on the size of the residency program and the number of procedures performed by a resident, a basic increase in speed and skill outside the operating room proves to be a beneficial aspect for both the patient and hospital.

Medical bills create an enormous burden on both the patient and society. It is apparent surgical trainees need as much practice as possible, but using the operating theater for hands-on training is not a cost effective method. The use of simulation for training avoids the already escalating cost of operating time. Moving training to simulators circumvents cost, but it is essential that the skill developed on simulators transfers to real scenarios inside the operating room.

2.4.4 The Need for Validity

The state of inadequate deliberate practice in surgical training has created an increased need for simulation. The apprenticeship model also lacks in immediate feedback, an indispensable component of deliberate practice, and the development of expertise. Valid immediate feedback is an essential prerequisite for improvement of performance. In many surgical situations, days, weeks, or months may pass until a patients' development confirms or rejects the performance of surgery. Instructors must devise and successfully implement educational variations that improve specific components of the surgical environment [88]. This is an area where simulators could provide real-time feedback and instruction outside the operating room.

There are issues that arise with respect to a simulator's ability to provide feedback. For instance, when is a simulator's feedback valid? Also, when is the simulator itself deemed valid? Although many aspects of surgical training show promise for simulation, opponents to surgical simulator training raise concerns regarding simulator validity. The validity of a training simulator is directly related to the transfer of skills into the operating room. Without direct evidence from the simulator, conclusions that success on a simulator translates to improved skill in the operating room seems baseless, whether justly so or not. A simulator without validation will quickly become obsolete due to lack of implementation into training programs.

Simulator effectiveness is measured in terms of skill transfer from simulation to the operating room. The extent to which this transfer occurs is a measure of the "validity" of the simulator. Categories of validity exist to provide further dissection and analysis of a simulator's strengths and weaknesses. Definitions of validity in medical simulation are adopted from the psychological testing standards developed by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education [89]. The most common categories of validity found in the medical simulation realm include face, content, construct, concurrent, and predictive validity. Face validity is the most basic type of validity and also the easiest to attain. Face validity is the simulator's resemblance to the real-life scenario. This type of validity is usually satisfied by expert opinion via questionnaires or surveys. Content validity assesses the degree that the simulator can measure the task or procedure. Testing a simulator for content validity is commonly a non-statistical procedure based on a checklist or expert opinion. Face and content validity are often subjective in nature and carry little consideration in the overall assessment of the simulator device in the eyes of the scientific community.

Construct validity is the extent to which the simulator measures the specific trait or traits that it is claimed to measure. The first desired test of construct validity is proving that the independent variable or variables alter what they are meant to alter [90]. In many simulation validity studies, the skill level of subjects is manipulated and variables in the simulator are observed. By showing a differentiation in the scores between experts and novices, the simulator creates and displays whether the measurements made by the simulator are correctly identifying the quantifiable aspects of surgical skill.

In the realm of validity, simulators are also assessed on their reliability, or consistency of measurements. This relates to the accuracy and sensitivity of all sensors and measurement tools within the simulator. It is critically important that the simulator is reliable and provides consistent accurate feedback for deliberate practice during skill development.

Concurrent validity is the extent the simulator agrees with the gold standard for the surgical task or procedure. In a surgical simulator setting, the simulator scores are compared with the Objective Structured Assessment of Technical Skill (OSATS) or against a rival simulator that has been rigorously validated. The most difficult category of validity to attain is predictive validity. This is the ability to use a simulator to predict the future performance in real-world procedures or scenarios.

Although all or most of the categories of validity are still explored and used in simulation literature, specific terms have grown obsolete by the most recent publication of *Standards for Educational and Psychological Testing* [91]. The categories of validity suggest distinctly different types of validity, when actually validity is a “unitary” concept. This advocates a simulator is never partially valid, it is a holistic concept. The current measure of validity turns its focus to the types of validity evidence. This gives rise to the concept that validity requires extensive evidence along different measures representing a simulator that proves being a valid method in target the intended purpose.

Without extensive research in the field of validity and its nomenclature, it is difficult for a residency program to decipher a simulator's usefulness prior to purchase. Currently, Endoscopic and Laparoscopic surgery lead the field in simulation and validation studies. In a comprehensive review of simulators, the European Association of Endoscopic Surgeons (EAES) established consensus guidelines for validation of virtual reality surgical simulators. In this 2005 report, the EAES systematically analyzed eight commercially available endoscopic and laparoscopic surgical simulators and awarded each with a "level of recommendation." The level of recommendation is based on a 1-4 scale (1-highest, 4-lowest), where specific validity criterion and studies, termed "level of evidence," must be achieved before a simulation attains a higher level of recommendation [92].

Simulator manufacturers face a daunting uphill battle before gaining wide acceptance in the medical community in the absence of rigorous validity studies. Standardization of guidelines for simulation validity is important to establish and adhere to across all disciplines. Surgical residency programs, already hesitant to widely implement simulation, need evidence-based studies showing transfer of skills in the operating room. These studies are necessary to gain acceptance in the medical industry comparable to the level of acceptance in the aviation industry.

2.4.5 The Spectrum of Medical Simulators

The range of medical simulators varies greatly, from tabletop simulators that do not resemble a human body to life-like mannequins that incorporate anything from an arterial pulse to blinking eyes. Again, it is imperative to realize that the appearance or physical fidelity of a simulator does not determine its validity in improving and transferring skills to the operating room. Validity is often times confused with fidelity.

However, fidelity in the field of simulation refers simply to the realism of the simulator. High fidelity simulation often creates an initial appeal to the user, but the validity of the simulation is not tightly correlated with the fidelity.

Medical simulators vary greatly on the fidelity spectrum. For example, box trainers are the lowest classification of physical fidelity. Although these are low-fidelity, they provide the benefit of being simplistic to operate and low-cost. Unfortunately, the very nature of box simulators is often misunderstood or misjudged, causing the simulators to be dismissed because of their simplicity.

The term “box” arose because trainers were initially designed with a simplistic box to obstruct the direct view of tools from the user, simulating the body cavity in a real surgery. Although the realism of the box is low, the box adds a similar amount of difficulty to the trainee. The first widely successful box trainer was MISTELS (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills). Inside the MISTELS box exists a camera to simulate the imaging technology used during an operation. The tools inside the box are the exact tools used during surgery, although the tasks involved are simplistic such as knot-tying and placing objects on pegs [27]. Although simplistic, the MISTELS system has shown construct validity [27, 93], improvement of skill [94], correlated in vivo surgery performance [29], and demonstrated reliability in skill measurement [28]. In addition to varying in fidelity, medical simulators also vary in terms of scope. Medical simulators fall into one of three categories: task, full procedure, and full body mannequins [31].

Task trainers are designed to enhance individual tasks within a procedure or underlying abilities needed to successfully complete a procedure. Examples may include picking up objects with surgical tools, completing suturing on inanimate objects, or familiarizing with a bone saw by sawing through an artificial bone. These task trainers are mainly focused on individualized skill and rely more on the psychomotor ability of surgeons and physicians.

The next step in complexity moves to full-procedure simulators. Full procedure simulators are typically more expensive than task trainers. These systems test the trainee along many procedural steps, and may also require a larger range of skills such as teamwork and communication. Full procedure simulators offer the benefit of testing trainees along all steps of a single procedure, such as a laparoscopic cholecystectomy. However, these simulators also have an inherit disadvantage: full-procedure simulators are not as flexible and versatile at changing the type of procedure. This requires a residency program to spend a large sum of money for a single procedure, or at most a small number of procedures.

The highest tier of complexity is the category of full-body mannequins. These advanced simulators focus on realism of the body and the procedure. These expensive simulators usually function on a wider range of procedures than full-procedure based simulators. For example, SimMan™ (Laerdal Medical) is capable of simulating a variety of simulation scenarios from cardiac failure to endoscopic intubation. These mannequins are highly advanced, which enables users to engage in near realistic training, but they also require a large investment from educational institutions. In addition to cost, these mannequins often require that another human controls the simulator while the trainee be in the simulation environment, which further increases the cost of utilizing the simulator.

Like any simulation, drawbacks always exist in simulating a real experience. A single medical simulator, in the foreseeable future, will never fully function as a human being. Simulators today are unable to incorporate the human aspect of emotion or the feedback of pain. This is less of an issue when simulators replicate a human under anesthesia, but is still an important omission. Simulators address the common situations of medical procedures, hoping to develop enough skill on the common situation so that the physician can address specialty or unique cases. When looking at the drawbacks, it is most important to keep an objective, balanced view of the evidence supporting the benefits of practice, including both risk and cost.

2.4.6 The Current State of Orthopaedic Simulators

The current state of orthopaedic surgical simulation lags behind other surgical disciplines, especially laparoscopic surgery. In the past two decades there has been very little research attempting to create an orthopaedic surgery simulator, and even fewer studies attempting to validate orthopaedic simulators. Due to orthopaedics being heavily dependent on technical skill, orthopaedic simulation holds great potential creating a heavy impact on improving surgical skill.

Developed orthopaedic simulators exist in two main categories: software-based simulations and orthopaedic simulators incorporating a haptic feedback device. In 2000, Sourin and Sourina developed an orthopaedic simulator for fracture fixation. This simulator was purely software-based with point-and-click methods for positioning screws in the virtual environment. This research focused on the development of software and did not include any experimental studies or validation procedures [95, 96]. Around the same time, Tsai, Hsieh, and Jou touted a software-based orthopaedic simulator for 3D manipulation of bones created from radiological data [97].

Blyth, Stott, and Anderson created a virtual environment for orthopaedic hip surgery, denoted the Bonedoc dynamic hip screw (DHS) simulator. This full procedure virtual simulator for DHS allows users to point-and-click different locations and tools used for the implant [98, 99]. Although these software-based simulations are good for procedural steps, they lack the technical skill input required to be a surgeon. Clicking and positioning screws on a computer screen with a mouse is vastly different from actually holding a drill in the hand, drilling real bone.

The invention and application of the PHANToM® haptic device (SensAble Technologies) allowed for the creation of haptic-feedback in the simulation. Photographs of PHANToM® implementation in orthopaedic simulators are shown in Figure 4. The current state of orthopaedic simulation has directed a large effort towards reproducing the appropriate haptic feedback encountered while drilling through bone with the

PHANToM® device. Tsai, Hsieh, and Jou, six years after developing their first orthopaedic simulator, attempted to model the forces and vibrations using a voxel-based model in the virtual theater [100]. Pettersson et al. performed a similar study advancing the haptic technology using CT data for automatic segmentation for use in a visual and haptic virtual environment [101]. Using the PHANToM® device for input to an orthopaedic hip drilling simulator (Simulutions TraumaVision VR), Froelich et al. found no significant difference in the wire accuracy between novice residents (PGY 1-2) and experienced residents (PGY 3-4). The less experienced residents used more fluoroscopy time and used more attempts to correctly place a guide wire in a virtual hip [1]. Vankipuram et al. used a haptic-feedback drilling model to virtually drill horizontally through the femoral body with accuracy analyzed with respect to virtually shown targets. The study shows differentiation of skill between novices, residents, and experts. By observing that the experts time to complete the exercise was longer than the novices and residents, an unexpected result, the results also noted that the experienced subjects were more accurate in drilling [2]. This review of orthopaedic simulators shows the current field for hip fractures is small. Unfortunately, even within the limited pool of simulators, only a select few out of the pool have attempted validation studies.

Figure 4. PHANToM® haptic devices shown in orthopaedic drilling simulation. From Tsai and Tsai (left) [100], Froelich et al. (middle) [1], and Vankipuram et al. (right) [2].

Other researchers agree that the orthopaedic field shows promise for improving guide wire placement through simulation [102]. The current state of orthopaedic simulation employs a haptic feedback device to interface a drilling procedure in the virtual environment. Haptic feedback is an important feature to simulate in the virtual environment and adds to the user learning. However, the haptic feedback devices have inherent limitations. Due to the frequency response time of the devices, sensations of hard virtual surfaces are difficult to simulate. This is a common problem with haptic feedback devices. The haptic feel of contact with hard surfaces, such as bone, tend to feel soft and spongy, leading to instability in the device [103-108]. Convex surfaces, such as the edges of bone, are difficult to recreate in a virtual environment, often giving the sensation of oiled glass or generally slippery surface [108]. Further complicating the problem is the act of using a haptic device for drilling. The response rate of the device is too slow to emulate the correct vibration while drilling, leading to a drilling vibration which can only emulate vibrations below 1 kHz [109].

Two possible options exist to advance haptic sensations during a bone-drilling task. First, technologies of haptic feedback devices could be advanced, so that current limitations of haptic devices such as hard, convex surfaces feel more realistic, and sampling rates are increased to adequate levels to account for drilling vibration. Secondly, new creative methods can be researched to decouple orthopaedic drilling simulation from haptic devices. In addition to new technology creation, more research needs to be completed in the area of validation. This lagging field points to building simulators that further advance common skills needed during orthopaedic surgery.

2.5 Targeting Hip Fracture Surgery Simulation

The previous section explored the existing state of simulation in the medical field, highlighting the components to construct a successful simulator. In addition, orthopaedics was identified as a field lagging behind the rest of the medical community in simulation. The following section focuses on explaining how hip fracture surgery simulation is an untapped field, which could make a drastic positive impact on the current state of health care. Before explaining the technology involved in constructing a hip fracture simulator, a brief background on the classification of hip fractures will be explored.

2.5.1 The Burden of Hip Fractures in the United States

Hip fractures are currently one of the most serious health problems afflicting the aging population [110]. It is estimated that 247,000 hip fractures occur yearly in the United States, with a majority occurring in the population over 45 years old [8]. The incidence of hip fracture is on the rise, with an expected 512,000 hip fractures per year by 2040. The cost of these fractures are also expected to rise from $7 billion per year [8], to nearly $16 billion per year by 2040 [111]. Each hip fracture is estimated at costing between $39,555 and $40,600 in the first year after surgery [112]. Hip fractures have the highest cost of any orthopaedic procedure after surgery, amounting to an extra $11,241 per year in extra health costs [9], with much of the cost passed on to society. Of course, the problem does not exist solely within the United States. As life expectancies in developed nations increase, so does the likelihood of hip fractures. Worldwide, it is projected that 6.26 million hip fractures will occur by 2050 [113].

Hip fracture surgery is one of the most common orthopaedic surgeries, behind only knee arthroscopy, shoulder arthroscopy, removal of a support implant, and total knee replacement [114]. Due to the severity of mobility loss, pneumonia, bedsores, and pain in many cases after hip fracture, dire consequences often arise. Due to these consequences, hip fractures are deemed more urgent than the more commonly encountered orthopaedic surgeries. Bannister et al. found mortality rate of 26 percent within three months of the trochanteric fracture and 35 percent within one year [115]. In a study of 223 patients, Wolfgang, Bryant, and O'Neill found similar results of 13 percent mortality rate within six months of surgery [116]. Estimates show after surgery, 15 to 25 percent remain in long-term care institutions for more than a year after surgery. Additionally, 25 to 30 percent rely on mechanical devices for mobility aid [8]. At a relatively young age of 50 years old, women have a lifetime hip fracture risk of 16 percent [117]. In fact, hip fractures generally increase exponentially with age. Exponential growth of hip fracture incidence is shown in Figure 5. The majority of fractures occur between the ages 75 and 84 years old. In 2005, an estimated 43.6 percent of the population 85 years old or older had a hip fracture. Using 2004 Medicare data, hip fractures have a mortality rate of 21.9 percent within 360 days of surgery [118].

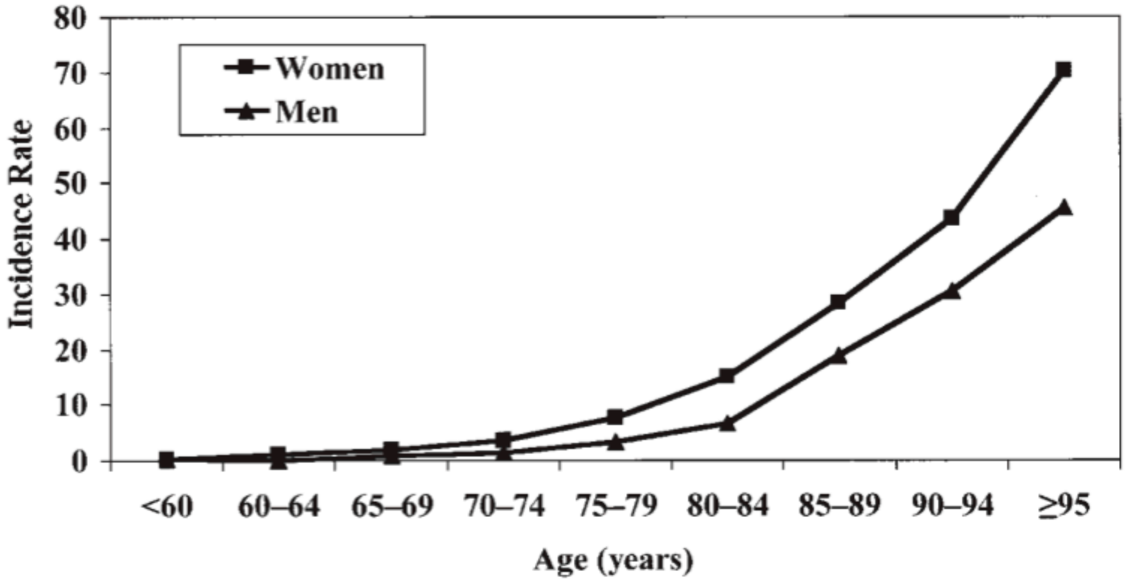

Figure 5. Age incidence rates (per 1000 person-years) of hip fracture, from [119].

These statistics show that hip fractures are a growing problem both in the United States and other populations throughout the world. Due the enormous expenses and extremely high mortality rates associated with hip fracture surgery, it is important that surgeons achieve the best possible outcome with every patient. Loss of mobility from poorly repaired fractures can lead to death. Before exploring the area of hip fracture simulator possibilities, it is important to understand the types and classifications of hip fractures.

2.5.2 Classifications of Hip Fractures

Hip fractures occurring in the femur exist in two broad categories: intracapsular and extracapsular. Intracapsular fractures occur above the intertrochanteric line, including the neck and head of the femur. Extracapsular fractures occur at or below the intertrochanteric line, including trochanteric and subtrochanteric fractures. The most common type of hip fracture surgery includes femoral neck (intracapsular) and peritrochanteric (extracapsular) fractures. Due to the frequency, classification systems will be explored for both femoral neck and intertrochanteric fractures. This enables a better understanding of the challenges surgeons face for a broader knowledge before designing a simulator.

Femoral Neck Fractures

In 1961, R.S. Garden proposed a four-stage classification system for fractures of the femoral neck. These fractures vary in severity and can result in disrupted blood supply to the femoral head. When blood flow is disrupted, avascular necrosis (AVN) results and the head of the femur dies. Stage I, the least severe, is an incomplete subcapital fracture. This incomplete break of the femur, called greenstick, gives a radiological illusion of the appearance of the impacted capital region. Stage II defines a complete subcapital fracture without displacement. The femoral neck is completely broken, but there is no rotation or tilting to the capital fragment. Stage III is again a complete subcapital fracture, but this type of fracture includes partial displacement of the capital fragment. Also, the reinacular attachments have not completely severed from the posterior surface of the neck. Stage IV consists of a complete subcapital fracture with full displacement of the capital fragment and reinacular attachments are severed [120]. Radiology images of the classifications are shown in Figure 6.

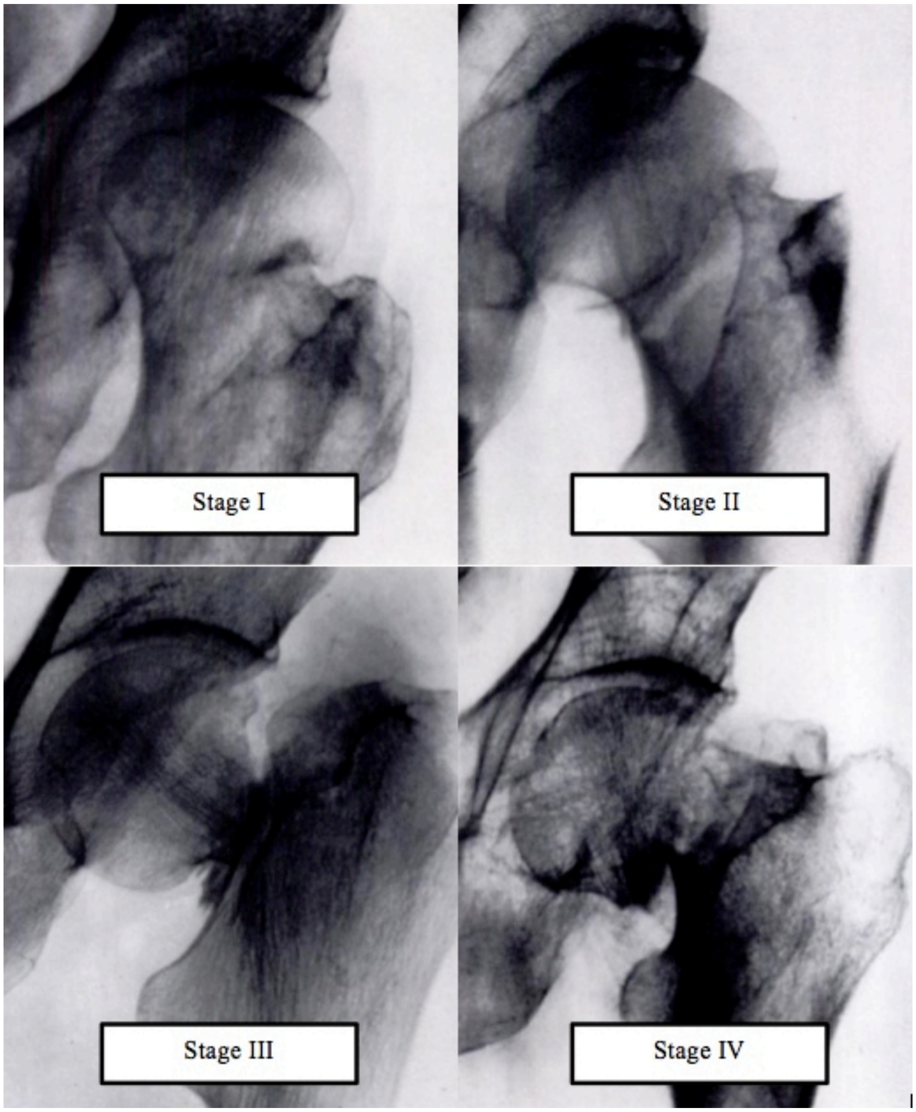

Figure 6. Garden’s classification of subcapital fractures. Images from [120].

Intertrochanteric Fractures

Similar to intracapsular fractures, a classification of trochanteric fractures was proposed by E. Mervyn Evans in 1949. In the literature, Evans classified trochanteric fractures as either Type 1 (stable) or Type 2 (unstable). The stable fractures were found to occur 72 percent of the time, with 90 percent of the stable fractures being fractured but undisplaced. The unstable fractures (Type 2), accounted for 28 percent of trochanteric fractures [121]. Jensen and Michaelsen further revised Evans’s classification system in 1975 to the classification commonly known today as Evans classification, less commonly but sometimes refereed to as the Evans/Jensen classification. Type I are defined as undisplaced, stable trochanteric fractures. Type II fractures are reclassified to encompass stable trochanteric fractures with displacement. Type III, IV, and V are created to represent unstable fractures. Type III fractures are classified as 3-part trochanteric fractures without posterolateral support and with separation of the greater trochanter. Type IV fractures included 3-part trochanteric fractures without medial support and separation of the lesser trochanter. Type V trochanteric fractures include 4-fragment fractures. These include loss of medial and posterolateral support and displacement of both the greater and lesser trochanter [122].

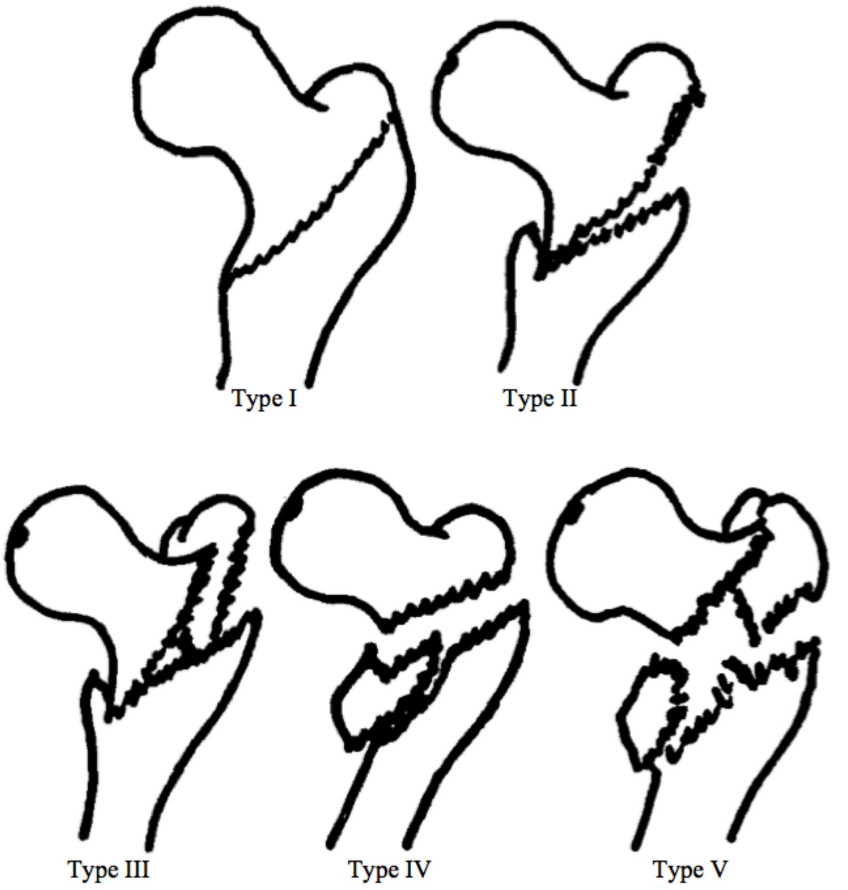

Figure 7. Evans classification of trochanteric fractures. Images adapted from [122].

2.5.3 Treatment of hip fractures

Like any other fracture, hip fractures differ in severity. One could argue that no two fractures are identical, and furthermore, that fractures with the same classification are sometimes treated differently due to comorbidities. Though fracture treatments are complex and require thought and planning, treatments can be narrowed into two main categories: arthroplasty and internal fixation. The most invasive hip fracture surgery is arthroplasty. This category can be separated into two subgroups, total hip replacement (total hip arthroplasty), where the femoral head and acetabulum of the pelvis are both replaced, and hemiarthroplasty, where commonly only the femoral head is replaced. Radiographic images of arthroplasty are shown in Figure 8.

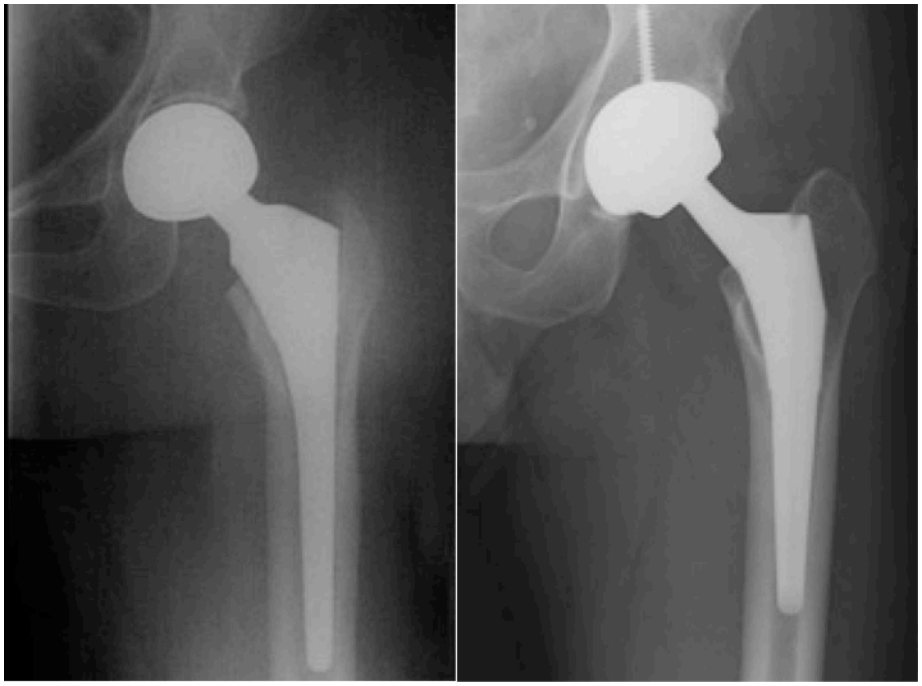

Figure 8. Hemiarthroplasty (left) and total hip arthroplasty (right), from [123].

The second method of repairing hip fractures is internal fixation. Internal fixation requires the use of plates, screws, and/or intramedullary devices to return the existing bone to a stabilized state and aligned position for healing. In consideration of femoral neck fractures, they exist in the Garden classification as either nondisplaced or displaced. In a nondisplaced fracture, femoral neck fractures are commonly managed using multiple stabilization screws from the lateral portion of the proximal femur to an anchored location within the femoral head [124]. A nondisplaced femoral neck fracture fixation is shown in Figure 9.

Figure 9. Undisplaced femoral neck fracture fixation using cannulated screws, from [123].

Repair methods for displaced femoral neck fractures depend on a variety of conditions, such as the patient's health, age, bone quality, and the surgeon's preference. Though maintaining the patient's own femoral head is preferred, if blood supply to the femoral head is disrupted, avascular necrosis (AVN) can emerge. This is more common in the case of a displaced femoral head. Avascular necrosis is the result of insufficient blood being supplied to the femoral head, which results in weakening, followed by a possible collapse of the femoral head. A radiographic image of AVN in the femoral head is shown in Figure 10. Internal fixation techniques often require operative revision in the form of partial or total arthroplasty [124].

Figure 10. Collapse of the femoral head due to avascular necrosis. From [125].

Pertrochanteric fractures, including both intertrochanteric and subtrochanteric fractures, are most often repaired using internal fixation techniques. Internal fixation of these fractures occurs in two separate categories, extramedullary and intramedullary. The most common extramedullary implant is the sliding hip screw, also called the dynamic hip screw (DHS) or compression hip screw (CHS). This device consists of one large lag screw driven through the lateral cortex of the proximal femur to the tip of the femoral head. This lag screw is attached to a plate on the lateral side of the femur with several additional screws attached to the upper femur. If needed, an additional screw can be driven to the femoral head for rotation stability. The term sliding, dynamic, or compression originate from the lag screw ability to slide along its axis inside the femur to impact firmly throughout the healing process, especially when a patient places a load on the hip, such as standing or walking [124].

Figure 11. Fracture fixation using a sliding hip screw, from [123].

Intramedullary internal fixation uses a similar idea to the sliding hip screw, using the implant to support most of the biomechanical load. Instead of a using a plate on the lateral cortex of the proximal femur, an intramedullary device, or intramedullary nail (IMN), is placed through the greater trochanter and down the middle canal of femur. In addition, an interlocking sliding lag screw is placed through the nail into the femoral head to promote impaction between the fragments. Often times the intramedullary nail is locked with a screw through the distal region of the femur. One advantage of this device over the sliding hip screw comes from the location of the device near the central axis of the femur so that biomechanical torques on the implant are much less than a plate attached to the outside of the femur. This device also comes with inherit risks, such as intraoperative fractures when driving the nail [124]. A radiographic image of an intramedullary nail for hip fracture fixation is shown in Figure 12.

Figure 12. Femoral fixation using intramedullary nail, from [123].

2.5.4 Fluoroscopic-Based Navigation Skill

Implementing simulation for hip fractures would have the greatest overall, worldwide impact on skill training due to the high occurrence of hip fractures throughout the world. Further narrowing the target for development of a hip fracture simulator, it is important to notice the common skills between all types of hip fracture procedures. Accurate placement of implants is an essential skill to all internal fixation procedures of the hip, and developing a simulator for a task used during multiple procedures will enable the greatest contribution to the orthopaedic field.

The current trend in orthopaedic surgery is pushing towards minimally invasive procedures. In hip surgery, minimally invasive procedures lead to better outcomes, less operating and hospital stay time, and less blood loss during the procedure [126]. However, minimally invasive surgeries require a high mental workload due to their complexity. Most minimally invasive hip surgeries are done using the aid of a fluoroscope. A fluoroscope is an imaging device that uses x-rays for 2D real-time static images of the internal structures of the patient. Using what is termed “fluoroscopic-based navigation,” surgeons are able to place implants and screws in accurate positions without direct sight of the implants inside the body. The fluoroscope is typically used only to show static images of an internal structure. Because live movement is absent from static images, a surgeon must use many different images throughout the procedure, with each image exposing the operating room to more x-ray radiation. In addition to complexity, the surgeon must be able to mentally map real-world movements to reflect accurate adjustments on the fluoroscopic image. In addition, the surgeon must be capable of operating under an elevated, complex workload due to the need to recall and image the intricacies of the human body in a 3D environment only using a static 2D image. Fluoroscopic-based navigation is an acquired skill that is remarkably time-consuming, even for advanced surgeons. For example, placing one distal locking screw into an intramedullary nail is found to take on average 108 seconds [127]. Another study found that placing screws for proximal and distal locking of an intramedullary nail took more than twice as long as driving the nail [128]. Fifty-two percent of all fluoroscopy time is attributed to proximal and distal locking in intramedullary nails [129]. In addition to the increased skill required to capture fluoroscopic images, advanced surgeons can even experience difficulty reading and interpreting the images. Noorden et al. found in hip fractures that nearly 8 percent of the time, implants protrude into the hip joint, potentially causing additional damage unknown to the operating surgeon [130].

The current solution to the complexity of fluoroscopic-based navigation is to introduce more technology into the operating room. This additional technology is called image-guided navigation (IGN) or computed-aided orthopaedic surgery (CAOS). These consist of overlays and additional technology used with the fluoroscopic images to direct and instruct the surgeon on placement of implants [131-137]. However, additional technology in the operating room does not address the underlying issue of developing the skill of fluoroscopic-based navigation in novice surgeons. The push towards minimally invasive surgery, coupled with the complex activities required for fluoroscopic-based navigation, makes for a strong argument in support of the development of a simulator that aims to develop surgical skill using a fluoroscope.

2.5.5 Fluoroscopic-Based Wire Navigation of the Hip

Interpreting fluoroscopy is a common skill used throughout hip fracture surgery, making it an ideal target for simulation. In addition to fluoroscopy, a universal psychomotor skill used in all hip fracture surgeries must be identified to use in conjunction with fluoroscopy skill development. The commonality among hip fracture surgery is the use of a surgical drill, whether it is used for reaming a hole, driving a lag screw, or locking an implant. Accurate drilling is an essential skill used in hip fracture repair. Although drilling is used in a variety of applications using a variety of hardware, most hip fracture surgeries utilize guide wires, commonly called Krischner wires, or simply K-wires, named after their inventor Martin Krischner. These thin metal wires are commonly driven into the femur as a preliminary task in a hip fractures surgery. Due to the small diameter of the wire, they can be drilled and repositioned. Although repositioning is not optimal, reposition a guide wire has reduced potential consequences compared to an error with a larger diameter drill. Guide wires are driven during many types of internal fixation procedures. In intramedullary nailing, a guide wire is initially drilled through the internal canal of the femur. Upon completion, the wire guides the reamer attachment for the drill down the canal by constraining its trajectory to the set trajectory of the wire. In nondisplaced femoral neck fractures, guide wires are used as the preliminary task before fixating the fracture with cannulated screws. Cannulated screws are hollow in the center, designed to fit over and around the guide wire for accurate placement inside the femoral head. Accuracy of guide wire placement is most crucial when inserting a sliding hip screw. The guide wire determines the precise placement of the sliding hip screw. Due to the weight-bearing purpose of the sliding hip screw, incorrect placement leads to failure of the implant. Incorrect placement of the sliding hip screw is the main cause of failure, with a failure rate of 10-16 percent [115, 116].

Guide wire accuracy is a skill used across many types of hip surgery. Due to this commonality, it is an ideal task to target for a hip fracture simulator. Specifically, due to the severe consequences resulting from inaccurate placement of the lag screw in the femoral head, guide wire placement plays an increasingly important role. Although guide wire placement is important among many types of hip fracture surgeries, the importance and large number of sliding hip screw procedures and intramedullary fractures makes it an exceptional starting task for the creation of a hip fracture simulator. Accurate placement of a guide wire in the femoral head could also lead to transferable placement accuracy to other screws needed during hip fractures.

2.5.6 Tip-Apex Distance

Lag screw placement accuracy for use in placing intramedullary and side-plate implants is fundamental to the success of surgery. Lag screws placed too far from the apex of the femoral head create the risk for mechanical failure and cutout. Lag screws placed too close to the apex of the femoral head risk penetrating the acetabulum, another undesired result. The surgeon faces complex decisions determining a sufficient placement for lag screw without risking penetration of the joint. To aid in this decision making process, the tip-apex distance (TAD) was developed for a quantitatively determining the sufficiency of a lag screw placed in the femoral head.

Early research was aimed at finding a placement location predictor and creating recommendations for hip fracture fixation. Failure of pertrochanteric fractures is also highly correlated with the positioning of the lag screw in the femoral head [138-141]. This is because the strength of bone-screw interface depends heavily on the location of the lag screw. That measure of strength depends on the number of cortices of fixation and the varying quality of bone, commonly a consequence of osteoporosis [138]. As such, it becomes important to drive the lag screw into the subchondral bone for rigid fixation, with early estimates recommending 11-25 millimeters from the subchondral cortex for non-telescoping implants [140]. Telescoping implants, such as the sliding hip screw, can be placed deeper into the femoral head because of the lower risk of penetration into the joint during healing. Deep insertion with the tip of the screw into the subchondral cortex gains maximum protection against deformity due to the increased strength of the subchondral bone [141]. It is also recommended that the implant be placed in a location where it would need to break through the maximum amount of bone to cutout.

The most common fixation failures are caused by cutout of the lag screw through the head of the femur. Baumgaertner et al. developed a simple universal method of describing the location of the lag screw within the femoral head for a sliding hip screw, called the tip-apex distance. With this measurement, the researchers have demonstrated that TAD is a strong predictor for cutout in pertrochanteric fractures [142]. Baumgaertner et al. defines the tip-apex distance as “the sum of the distance, in millimeters, from the tip of the lag screw to the apex of the femoral head, as measured on an anteroposterior radiograph and that distance as measured on a lateral radiograph, after correction has been made for magnification” [142]. The formula is presented in equation (1). The apex of the femur is identified as the projection of the midpoint of the neck parallel to the neck angle. The magnification of the radiograph is accounted for by dividing the known diameter of the lag screw by the diameter of the lag screw shown on the radiograph.

$$\mathrm{TAD} = \left(\mathrm{X}_{\mathrm{AP}} \times \frac{\mathrm{D}_{\mathrm{True}}}{\mathrm{D}_{\mathrm{AP}}}\right) + \left(\mathrm{X}_{\mathrm{Lat}} \times \frac{\mathrm{D}_{\mathrm{True}}}{\mathrm{D}_{\mathrm{Lat}}}\right) \qquad (1)$$

In the 198 fractures studied, no lag screw cut-out with a TAD of less than 25 millimeters, whereas 27 percent cut out with a TAD of more than 30 millimeters [142]. Using this data, a TAD goal of under 25 millimeters was established by surgeons for the placement of the lag screw. Explaining and familiarizing surgeons with the TAD measurement and its consequences has been found to have a significant impact on improving hip implant failure in pertrochanteric fractures [143]. While Baumgaertner et al. focused on researching the TAD in sliding hip screws, Geller found similar cutout findings with intramedullary implants. Of the 82 fractures studied, no cutout occurred with a tip-apex distance less than 25 millimeters. Shockingly, 44 percent of patients experienced cutout with a tip-apex distance above 25 millimeters [144]. This evidence shows that the TAD may be more critical for intramedullary devices.

Most surgeries are difficult to assess purely on quantitative measures. Aside from each patient's unique anatomy, variations also differ by surgeons in treatment methods and personal judgment. Many measures of surgical proficiency are often subjective and difficult to measure without an on-looking faculty member, as seen on the most commonly used assessment device, the OSATS. Unlike OSATS, the TAD is one concrete, proven quantitative measurement serving as a predictor of lag screws placement success in the proximal femur. Lag screw accuracy can foreseeably be automatically measured in the femoral head through technology implementation. In addition, this measurement could be easily implemented into a hip fracture simulator.

2.6 Summary

This chapter highlights how orthopaedics lags behind other surgical disciplines, such as laparoscopy, in the implementation of simulation training. The need for increased practice while reducing errors and cost argues shows the need for surgical simulation development in the future. To promote skill acquisition, these developed simulators need to focus on validation studies to ensure the simulation improves surgical skill. One largely undeveloped simulation field is orthopaedic surgical drilling in hip fracture surgery. The importance of wire placement in hip fracture surgery shows fluoroscopic-based wire navigation is an ideal candidate for simulation. The following chapter documents the creation of an effective wire navigation simulator through the use of augmented reality.

CHAPTER 3 – AIM ONE: CREATING AN AUGMENTED REALITY WIRE NAVIGATION SIMULATOR

3.1 Introduction

Apprenticeship training would greatly benefit from the addition of surgical training simulation. The need to reduce risk, errors, and cost in the operating room makes simulation a prime resource for supplemental training, without vastly increasing the already demanding work hours of surgical training residents. The field of orthopaedics, an intensely skill demanding surgical field, is lagging behind many other medical fields in the adoption of simulation. In particular, hip fracture surgery is a ripe candidate for simulation development due to the seriousness of the hip fracture surgeries in terms of skill requirement, mortality, and cost. In addition, the high incidence of hip fractures worldwide makes improvement of hip fracture surgery an area for arguably the largest potential global impact in the medical simulation field. Unique aspects present themselves between each hip fracture surgery; however, fluoroscopic-based wire navigation is ubiquitous in procedures across many different implant devices and fracture classifications. Fluoroscopic-based navigation demands considerable skill. This broad applicability makes fluoroscopic-based wire navigation of hip fractures an ideal target for development of an orthopaedic training simulator. The development of a fluoroscopic-based wire navigation simulator will enable a shift in the orthopaedic training curricula, in addition to giving residents a more efficient way for skill development through increased deliberate practice.

The following chapter shows augmented reality can be effectively used to develop an orthopaedic surgical simulator. The following development first explores design considerations in the pursuit of creating a fluoroscopic-based wire navigation training simulator. This research makes a case for aiding skill transfer by replicating specific stimulus-response cues present in the surgical drilling task, allowing for efficient transfer of skill from the simulator to the operating room. First, focused design specifications are determined through the identification and implementation of essential elements of the wire navigation task. Secondly, these design specification will be used to select specific hardware and software components for the simulator. Next, setup and calibration procedures will be outlined and clarified, along with a detailed explanation of the tool calibration algorithm. Then, the simulator's simplicity of operation will be explained, along with more advanced features embedded in the software. Finally, an overview of data collection and feedback capabilities will be discussed.

3.2 Design Considerations

3.2.1 Introduction

As discussed in Chapter 2, the essence of training simulator development is to advance skill transfer to a real scenario. Most simulators pursue realism by attempting to recreate physical aspects and design features. Even if cost and personnel were removed as limits to simulator development, a persistent demand exists for surgical simulators capable of duplicating or recreating inherently unique and complex aspects or systems within the human body. In addition, true assessment of skill transfer from the training simulator to a live surgical scenario is difficult because of the consequences and risks inherent in allowing inexperienced trainees to operate on live patients.

This design approach considers major sensory cues of the real fluoroscopic-based wire navigation task, and then concentrates on selecting and reproducing these cues. This approach filters out aspects of the wire navigation task extraneous to skill development. Although the selection and inclusion of wire navigation features in the simulator is initially subjective and theoretically falsifiable by future studies, the research provides a base for iterative incremental improvement.

3.2.2 Realism and the Fidelity of Simulation

The development of simulators requires recreating aspects of reality. The degree of apparent realism to which reality is recreated affects the cost, time, complexity, and technology required in development. The current trend in simulation is to strive towards building the most realistic simulator possible. These ultra-realistic simulators try to mimic minute and occasionally irrelevant details. These design decisions can require millions of dollars and a team of specialists to develop and maintain. This common trend in simulation idolizes physical fidelity as the epitome of simulator quality. Through the advancement of simulation, the term fidelity has erroneously been synonymized with physical realism. The misused definition of fidelity refers to merely the physical aspects of the simulator, pertaining to characteristics such as the realism of the graphics and resolution of projectors in the virtual environment. Establishing a compromise in terminology, the physical characteristics of the simulator are referred to as the "physical fidelity" of the simulator. The counterpart to physical fidelity exists as the psychological fidelity of the simulator. J.K. Caird defines psychological fidelity as "the degree that a simulation produces the sensory and cognitive processes within the trainee as they might occur in operational theaters" [145]. It is important to note that physical fidelity and psychological fidelity are not mutually exclusive. In some cases, physically fidelity has contributory effects to psychological fidelity. In other cases, higher physical fidelity simulators produce little to no quantifiable benefit in training over lower physical fidelity simulators [146, 147].

Transfer of skill is much more closely tied to psychological fidelity than the physical fidelity of a simulator. Transfer of skill from the simulator to real scenarios should be the pinnacle of all training simulator development goals. Surgical simulation has a particularly high need for psychological fidelity due to the intricacies of the human body, while high physical fidelity is unattainable and furthermore, irrelevant, in many scenarios involving organs, skin, and bodily fluid.

Although more difficult than assessing the physical fidelity of a system, the most important aspect of creating a training simulator is identifying and selecting the subset of aspects within the total tasks that need to be supported in the simulated environment. For instance, is displaying the accurate color of the rock lying on the ground in a virtual environment important? The answer depends on the type of virtual environment and the goal of simulation. Most would agree, in a virtual environment designed for military training, little effort should be spent on achieving the correct color shade of a rock. However, in a geology training simulator for identifying minerals, the correct shade and visual properties of the rock may be extremely important. This demonstrates when measuring skill transfer, some simulation details are irrelevant for some skills, but essential for others.

The assumption made in any training simulator is the existence of some similarity between the task in the simulated environment and the real world. This assumes that if similarity exists, training on a simulator will transfer to the real task environment. The challenge remains in identifying the similar elements that invoke the correct sensory and cognitive processes (i.e. psychological fidelity). Correctly distinguishing these details allows for optimization of the learning process. The developmental success of a training simulator depends on correctly differentiating aspects of a task that are critical for inclusion in the simulator from those which are irrelevant.

Lintern applies the term “relevant informational invariant” to describe characteristics which remain unchanged between two tasks that promote the transfer of skill [148]. It becomes theoretically possible to create a simulator with just the relevant informational invariants without adding extra, arguably frivolous physical fidelity to the simulator. Box trainers, perhaps by accident, possess many relevant informational invariants without high physical fidelity. Box trainers, which appear simplistic at first glance, have shown substantial success in the medical field because of their high transfer of skill to the real world. For example, box trainers containing only a few physical components from actual surgery have proven to be the most widely accepted and validated. An example is the MISTELS simulator [27-29]. The MISTELS only uses trochars from actual surgery, but abstractly identifies tasks such as pattern cutting and knot-tying as relevant tasks for transferability into the operating room. These low-fidelity devices contain lower face validity than other devices, but excel in identifying and requiring underlying skills in surgery, leading to strong overall validity of the simulator.

In the orthopaedic field, the current realm and field of simulators is nearly completely unexplored. Thus, it is difficult to build on previously established successful simulators. This work attempts to assess which aspects of wire navigation are important to recreate in a simulator and which can be ignored due to little or no influence on skill development and transferability.

3.2.3 Technique for Wire Navigation of the Proximal Femur

The first step in simulator development is understanding the precise technique involved in the simulation. This project emphasizes the fluoroscopic-based wire navigation task for placing a guide wire in the repair of an intertrochanteric fracture of the femur. Dissecting the fluoroscopic-based wire navigation task into basic sensory stimuli is one method for identification of important elements to recreate in a simulation. Physiologically, the surgeon, upon receiving the stimuli, responds with a cognitive decision-making process. The sensory information is processed by the visual, somatosensory, and auditory systems.

The wire navigation task begins after an incision is made on the lateral surface of the hip. Assistants spread the incision using retractors to give the surgeon access to the bone. Using the guide wire and the drill, the surgeon will make an initial guess for the correct drilling location and trajectory in the lateral surface of the femur. The surgeon uses static fluoroscopic images from AP and/or lateral views to assess the accuracy of the guide wire (Figure 13). From these viewpoints, the surgeon interprets and determines the relative three-dimensional placement of the guide wire in the proximal femur.

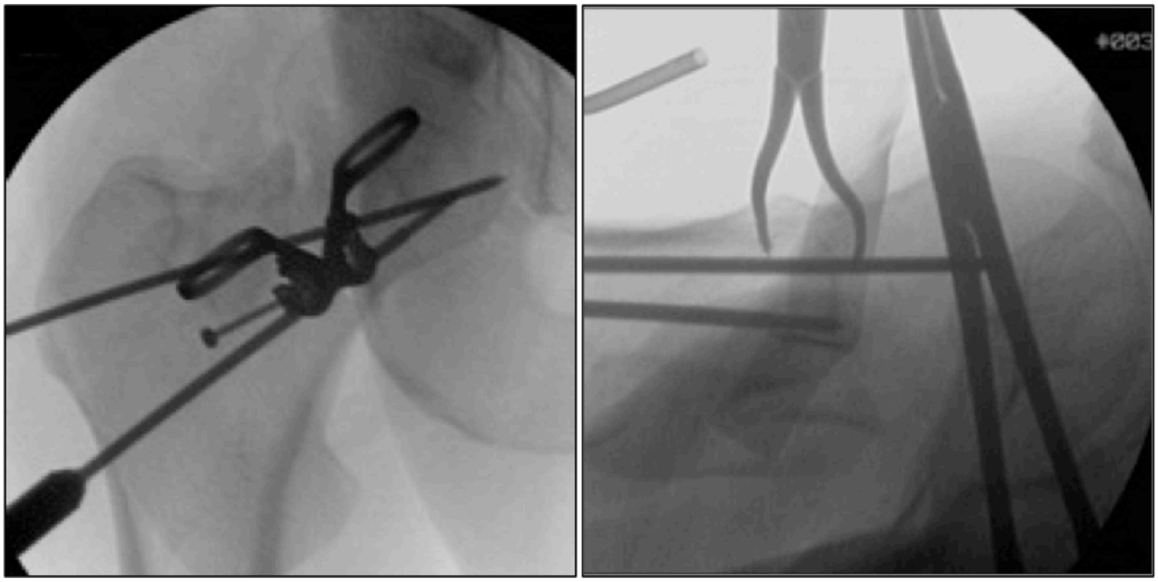

Figure 13. AP (left) and lateral (right) fluoroscopic images with guide wires during a hip fracture procedure.

From these 2D views, the surgeon is able to extract the spatial data to reconstruct a mental 3D anatomical model of the guide wire. The imagery serves as the feedback in the loop for assessing the accuracy of the guide wire. Once a satisfactory starting location and trajectory are achieved, the surgeon begins drilling into the femur towards the apex of the femur. The surgeon will intermittently check the progress of drilling using more fluoroscopic images. If the surgeon realizes the wire is not progressing accurately, the guide wire will be reversed and pulled out of the femur. The guide wire navigation process will then restart by finding a new entry location and trajectory. Once the surgeon accurately progresses towards the femoral apex, additional images will be used to ensure a TAD of less than 25 mm. Once the surgeon has reached satisfactory TAD without penetrating the acetabulum, the guide wire navigation ends and the succeeding process of inserting the lag screw ensues. A flow chart of guide wire navigation process is shown in Figure 14.

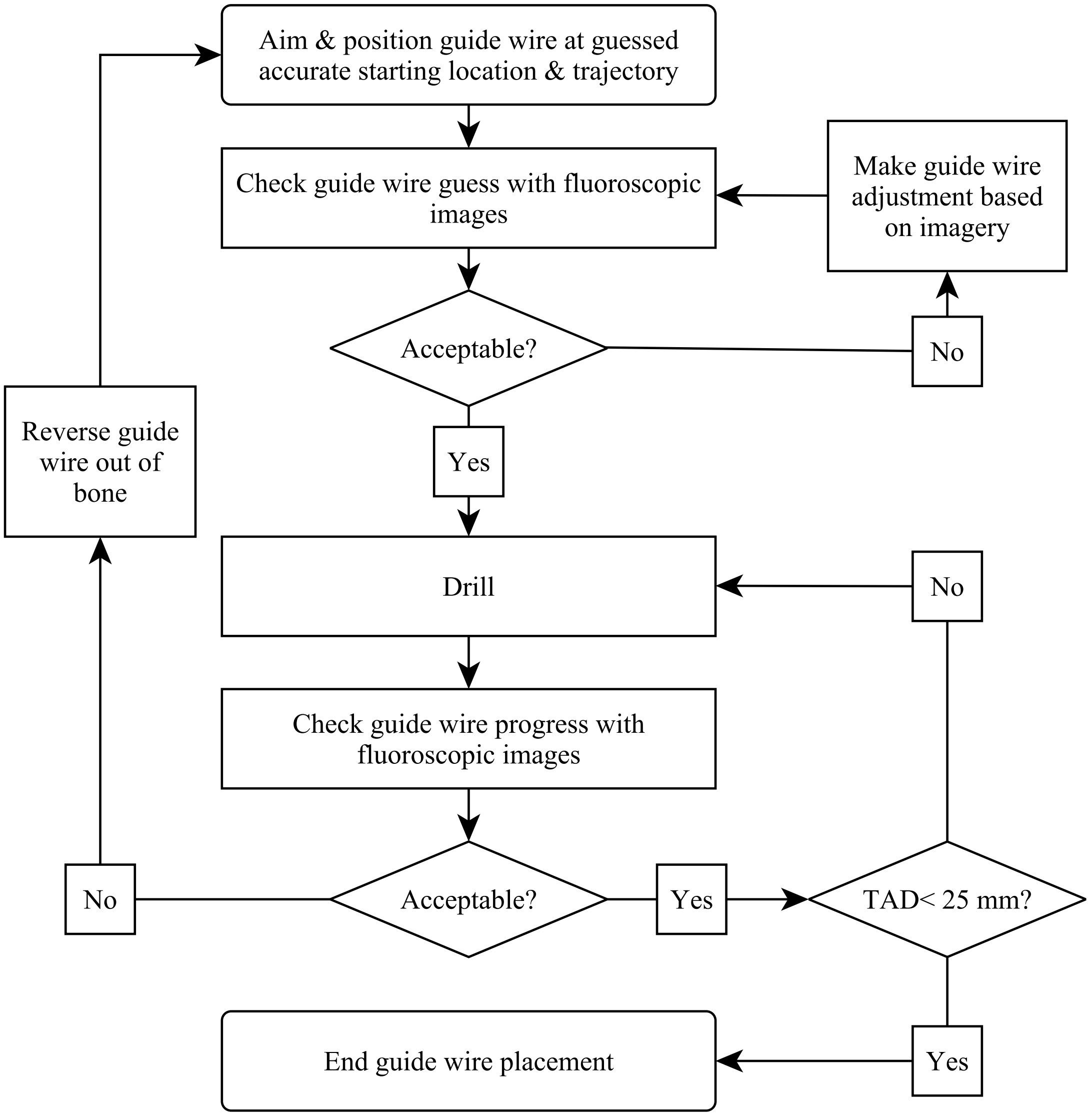

Figure 14. Flowchart for inserting a guide wire for intertrochanteric hip fracture repair.

3.2.4 Selecting Components to Recreate in the Wire Navigation Task

Through dissection of the wire navigation process, it is apparent the difficulty lies in the cognitive processes involved in coordinating drill and guide wire movements with the feedback imagery given through fluoroscopy. The surgeon must possess the skill to aim the wire towards the center of the femoral head using orthogonal views. The surgeon also must possess the somatosensory and visual ability to assess distances on the fluoroscopic images for both progressing the wire and assessing the TAD of the result. The visual, somatosensory, and auditory sensory cues will be used to select components for simulator recreation.

Visual Sensory System

The visual sensory system plays an integral part in the fluoroscopic-based wire navigation task. During the majority of the drilling task, the surgeon is either looking at the drill, around the sight of the incision, or viewing fluoroscopic images for trajectory and overall wire placement. In terms of simulation, when addressing visual aspects in the proximity of the drill site, how realistic does the hip region need to be? For this specific wire navigation task, adding realistic soft tissue and bodily fluid around the model merely adds an unneeded expense to the simulator for minor benefit. The most relevant visual aspect of the hip region is obstruction of a direct view of the wire inside the bone from the trainee. By obstructing the view, the surgeon must rely on the 2D fluoroscopic images to navigate the wire in the proximal femur. A simple covering over the drilling region fulfills the necessary objective.

When the surgeon is not viewing the drilling site, he or she is primarily viewing fluoroscopic images for a visual reference of guide wire location inside the proximal femur. The key aspect or goal of these static x-ray images is to accurately identify the location of the guide wire with respect to the key features of the proximal femur region, which includes the lateral cortex, greater trochanter, lesser trochanter, femoral neck, and femoral head. One of the main difficulties of this operating room task is mentally constructing and interpreting a 3D anatomical model of the proximal femur and guide wire from only 2D static image references. Therefore, view angles of the proximal femur in the simulator must be similar to the angles used during real surgery, commonly anteroposterior (AP) and lateral views, so that learned mental rotations of the 2D images to construct 3D mental models are the same as during live surgery.

It is important to acknowledge the quality of fluoroscopic images. Fluoroscopic images vary in quality based on fluoroscope settings and anatomical conditions. In simulator development, it is typically easier to replicate the proximal femur with sharp, crisp edges, and to avoid visual interactions between tissue and bone. Although visual conditions portrayed in a simulator are better than a real scenario, it may provide a benefit to the novice trainee that is still working on the cognitive task of constructing a three-dimensional mental model from two-dimensional images. This simulator eliminates image difficulties of a fluoroscope so the trainee can concentrate on the psychomotor task of accurate drilling.

Somatosensory System

The somatosensory system receptors are responsible for everything from pain to sensing temperature. The most important aspects of the somatosensory system for this specific motor task include touch and proprioception. The somatosensory system must be investigated at the macro level, in regards to body position and orientation, before investigating the more detailed region around the drill. The body position of the trainee must be similar to the position of the surgeon in actual surgery, ensuring the angles of the arm, wrist, and fingers are all oriented relatively the same. Although surgical drills vary slightly, to replicate the stimuli, the user’s hand must be oriented in a particular manner and supporting a weight similar to the actual surgical drill. The surgeon must position and angle the guide wire against the femur before drilling, making this action desirable for the replication of the haptic sensation of touching the lateral cortex of the proximal femur with the guide wire. This ensures the trainee has the possibility, just as in real surgery, of slipping or skiving off the bone before commencing the drilling procedure. When the drill is advanced, vibration exists, simulating the electric motor of the drill. In addition to the vibration while drilling, the proper approximate relationship of applied pressure to drill advancement must exist in the simulator. Although patients' bones vary in terms of hardness and density, if the coupling between pressure and drill movement is grossly incorrect, the trainee will advance the guide wire too fast or too slow when transitioned to real surgery. Lastly, drilling into the joint is a critical error that may occur during guide wire placement. Due to the severity of this error, it is valuable to replicate the haptic sensation of passing the guide wire out of the femoral head.

Auditory Sensory Systems

During the wire navigation task, the auditory system receives stimuli from both the background noise of the operating room and the sound of the drill. The noise of an orthopaedic operating room averages 66 dB [149], approximately the level of a normal conversation. A choice is made to omit the replication of background noise from the simulator because levels are comparable to a normal setting of simulation, such as a surgical training laboratory. The sound of the drill could provide supplementary information to an advanced surgeon. Of course, the accuracy of the drilling cannot be evaluated solely through auditory signals, but supplementary information exists to a degree in the auditory feedback. For example, a surgeon may decipher that the drill battery is low, or possibly gauge the hardness of the drilled material from to the drill sound. In a related technical field, advanced machinists are aware of tool wear and cutting conditions by the mere sound of the drilling tool. Similarly, the possibility exists that advanced surgeons possess the ability to collect tissue density data using the sound of the drill. Nonetheless, although the auditory sensory system plays a role in the wire navigation task, the most important aspects of the task still reside in the visual and somatosensory systems.

3.2.5 Comparison with Current Orthopaedic Drilling Simulators

After identifying a custom list of important wire navigation elements to include in simulation, it is interesting to reflect on previously developed simulators within the realm of wire navigation of the proximal femur. Analyzing the Bonedoc DHS simulator, utilizing a mouse for simulator input, reveals that only the visual system sector are developed in the simulator [98, 99]. The Bonedoc DHS completely omits all somatosensory aspects.

In simulators using haptic feedback devices [1, 2, 100, 101], the visual sensory system elements are incorporated into the designs. It is undistinguishable which existing simulators incorporate vibration into the simulation model. Even if vibration is included, due to the technology limitations of response frequency, the haptic feedback device of many of the higher frequency vibrational drilling cues would be excluded from the simulator development. In addition, the haptic feedback devices do not include the weight of a surgical drill in the hand of the trainee. Also, the haptic sensation of the hard cortical bone of the femur cannot be replicated with these haptic feedback devices due to additional technological response limitations. Although the simulators using haptic feedback address the somatosensory system in development, inaccuracies and technological limitations prohibit accurately recreating the haptic stimuli in the wire navigation task.

3.3 Design of the Simulator

3.3.1 Introduction

The primary goal of the orthopaedic fluoroscopic-based wire navigation simulator is to improve underlying skills integral to fluoroscopic-based wire navigation of the proximal femur. This skill is essential for internal fixation of hip fractures this technology could potentially improve drilling skill in other orthopaedic surgeries. The simulator implements all of the identified important stimulus-response cues of the wire navigation task while employing a novel approach, using a real drill and guide wire. With the direct view of the proximal femur obstructed, the participant is required to rely on virtually generated fluoroscopy to navigate the wire through the femur. The following sections detail the underlying design decisions of the current augmented reality fluoroscopic-based wire navigation simulator for hip fracture surgery.

3.3.2 Implementing Somatosensory and Auditory Components in the Wire Navigation Simulator

The most direct approach to incorporating the identified somatosensory and auditory stimulus-response cues is to use a real drill and wire, drilling through a replicated artificial proximal femur. Using a real drill for drilling incorporates realistic weight, vibration, and sound, accomplishing two development goals. The artificial femur ensures the wire advancement has a force and torque response corresponding to a realistic scenario. This potentially allows for replication of the perceived haptic sensations when the guide wire encounters the cortical region of the proximal femur. The use of a real drill with a replicate artificial femur fulfills all the somatosensory and auditory goals for the simulator.

Using a real drill simplifies the need to recreate force sensations while drilling but also adds complexity. For correlating visual stimuli with the somatosensory stimuli, it is vital to know the position of the guide wire with respect to the artificial bone throughout the simulation procedure. With a haptic feedback device, the position of the input angles and the graphical feedback are relatively easy to synchronize. With a real drill, the position must be calculated first then provided to the graphics system. These two steps must happen quickly in order for the graphics to be synchronized with the drill position.

To synchronize real-world movements with the virtual graphics component this simulator uses an electromagnetic motion tracking system. These systems use orthogonal coils inside a sensor to detect the strength of received electromagnetic waves, emitted from a transmitter. Using the signal strengths, six degrees of freedom are solved for each sensor, giving positional data and orientation data. The sensors are relatively small and do not require direct line of sight as compared to an optical motion system. This makes electromagnetic tracking devices less difficult to set up in operating rooms for procedures where the tracking sensors may be concealed from direct view. Electromagnetic sensors also require less space inside the tracking environment. One disadvantage of electromagnetic tracking is the common electromagnetic wave distortion encountered caused by ferrous metals in the environment, which reduces the device's positional accuracy.

3.3.3 Implementing Visual Components in the Wire Navigation Simulator

In addition to the somatosensory components, it is essential to replicate the visual sensory-response cues of the wire navigation task. Cylindrical foam covering secured with straps over the artificial proximal femur obstructs the direct view of the proximal femur from the trainee. The thickness of foam reflects the approximate thickness between the skin and lateral femoral cortex in an average patient. A semi-transparent 3D virtual model rendered with 3D graphics on a computer monitor provides the AP and lateral fluoroscopic views, from angles similar to those used in surgery. The challenging facet of the visual sensory system is the correct positional display of the guide wire as the trainee drills through bone. The electromagnetic tracking sensors collect spatial data for the guide wire as the trainee drills through bone. The foam covering, virtual 3D femur model, and electromagnetic positioning of the guide wire facilitate the visual recreation of the wire navigation task.

3.3.4 Augmented Reality

Through the implementation of the identified sensory components, the development naturally transformed to the realm of augmented reality in the reality-virtuality continuum. Milgram et al. first proposed the idea of this continuum, ranging from the real environment to the synthetic virtual environment. Between the two extremes exists a mixed reality (MR), containing both subcategories of augmented reality (AR) and augmented virtuality (AV) [150]. Augmented reality is a situation where a real world scenario is augmented by virtual reality, whereas an augmented virtuality environment is a virtual environment augmented by real world objects. This simulator uses real world objects, consisting of the real drill, guide wire, and artificial bone, and augments reality by virtually generating surrogate fluoroscopic images. This development benefits from real world drilling components and from radiation-free, virtually generated fluoroscopy.

3.3.5 Simulator Overview

The orthopaedic wire navigation simulator aims utilizes real world aspects of the wire navigation procedure while employing position tracking and three-dimensional graphics to simulate fluoroscopy. Using virtual fluoroscopy has a two-tier advantage over fluoroscopy in the real world. First, the trainee is allowed to practice as long as desired without the risk of overexposure to radiation. Second, the trainee can practice without the expense and inconvenience of an actual fluoroscope present. Training programs do not commonly have exclusive access to fluoroscopes for practice. Scheduling practice time on a fluoroscope involves working around the higher priority of actual surgeries in operating room. Even when fluoroscopes are available, additional costs are incurred because a radiology technician must be present to set up and operate the fluoroscope.

Generating the virtual fluoroscopy images for wire navigation requires tracking the position of drill and bone for calculating the relative position between the guide wire and the bone. By obtaining the relative position, a virtual image can be generated to reproduce a geometrically accurate pseudo-fluoroscopic image. The current system uses small, unobtrusive electromagnetic sensors that require less setup and space than commercial optical motion capture systems. While the electromagnetic system offers several advantages, it is necessary to avoid metallic interference to assure accuracy. It is important to note the simulator incorporates a tracking system with a modular integration. Although the system currently uses electromagnetic tracking, the setup can be easily swapped with another tracking device without disrupting the software. Accurate image generation enables a simulator capable of simulating fluoroscopy for use in training and practice of fluoroscopic-based wire navigation techniques.

Instead of using a haptic feedback device, the simulator uses a battery operated portable drill and a surrogate femur for actual drilling experience in the simulator. Another advantage of the design includes the capability to easily switch out the model and brand of drill without additional software augmentation. Incorporating an actual drilling experience with the augmentation of virtual fluoroscopy represents a novel implementation for orthopaedic simulation.

3.3.6 Hardware Overview

The hardware consists of four major components: a personal computer (3.06 GHz Intel Core 2 Duo MacBook Pro, 8 GB of RAM, running Microsoft Windows 7), an artificial large left femur (Sawbones Part No. 1129), a battery-powered portable electrical drill equipped with a guide wire (2.5 mm diameter), and Ascension 3D Guidance trakSTAR™ 6-DOF electromagnetic tracking sensors. Figure 15 displays the main components of the simulator.

The personal computer serves several purposes. First, the participant views the virtually generated fluoroscopic images on the display of the computer. The computer also serves as the input hub for both the user input and electromagnetic tracking output. The user commands the computer to generate a new fluoroscopic image using a mouse click. Upon clicking the mouse, the computer gathers sensor information and generates geometrically accurate fluoroscopic images.

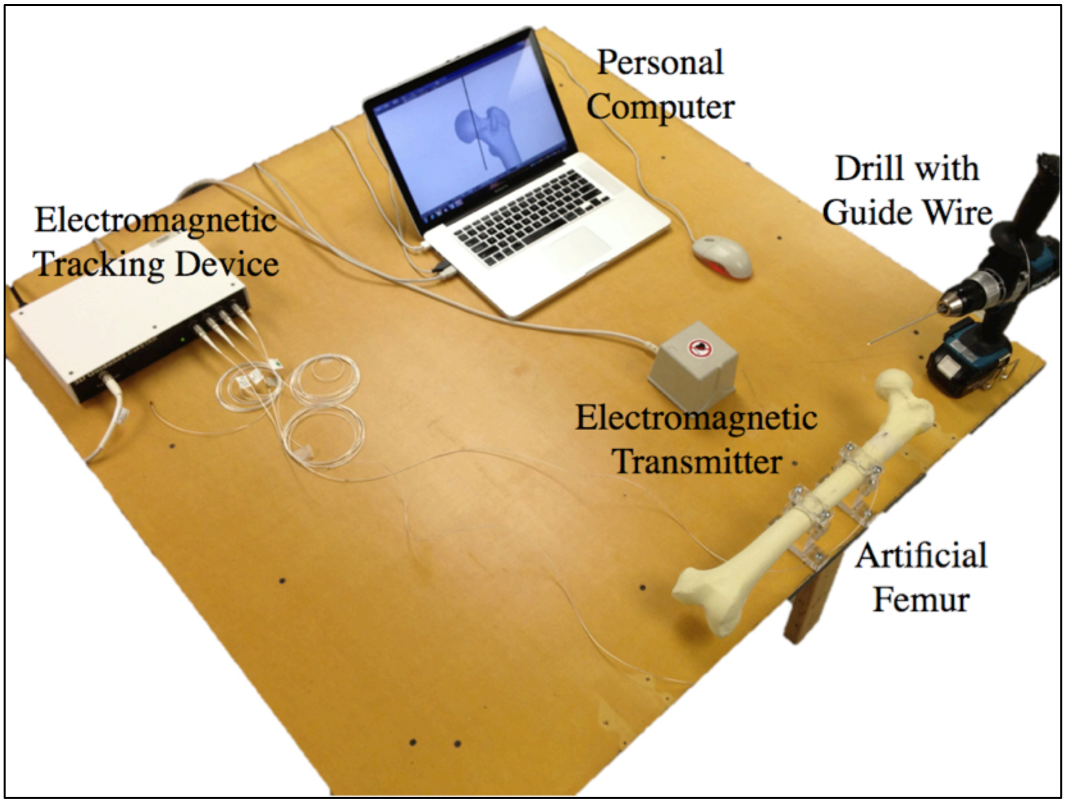

Figure 15. Main hardware components of orthopaedic augmented-reality wire navigation simulator.

The artificial femur serves as an expendable component of the simulator due to each trial creating a small hole in the proximal femur. They are held to the simulator table by two custom-made CNC machined clamps made from polycarbonate, shown in Figure 16. The most basic version of the femur produced by Sawbones® is priced at $12.25 per femur. This femur is a solid uniform material throughout. For heterogeneous bone density, Sawbones® produces a solid cortical artificial femur with a cancellous interior at a price of $14.25 per femur. Although artificial bones are expendable and are a reoccurring cost to the simulator, femurs have been successfully used through four trials without replacement. In addition, drilled holes can be filled in with similar density substances, such as wood glue, after several trials. This technique extends the life of each artificial bone, thus driving the cost of simulation to under $1.00 per trial.

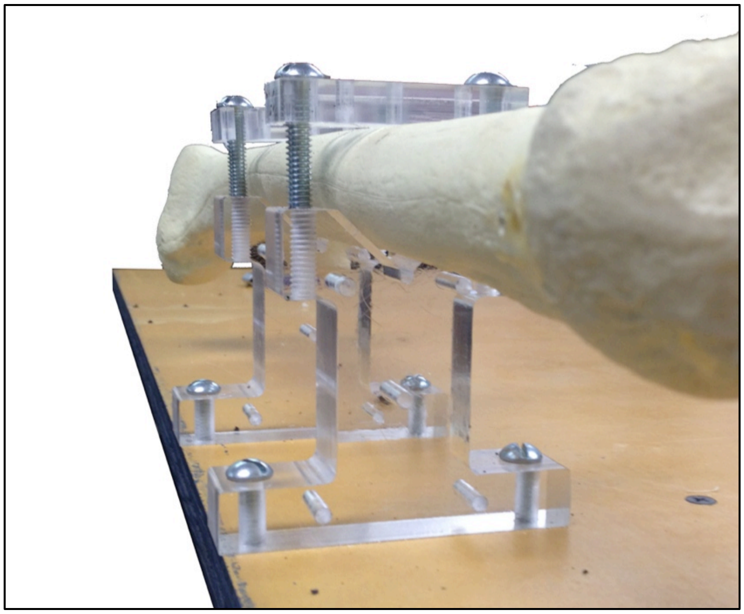

Figure 16. Custom clamps for securing artificial femurs.

The current electromagnetic tracking device, the Ascension trakSTAR™, can be expanded to track 16 sensors simultaneously, although the simulator only requires two sensors. One of the simulator's sensors attaches directly to the drill, tracking the movement and position of the guide wire. The other sensor attaches to the artificial bone, tracking the location in respect to the drill. The very small size of the sensors, under a centimeter in length, allows them to be nearly unnoticeable. The respective size of the sensor is shown in Figure 17. The current setup of the artificial femur attaches rigidly to the simulator table. Due to this rigid attachment, the sensor can be placed on any surface that does not move in respect to the artificial bone. In this situation, the sensor can be attached to the table enabling quick changing of the artificial femurs, allowing the subject to run through more trials in a given practice session.

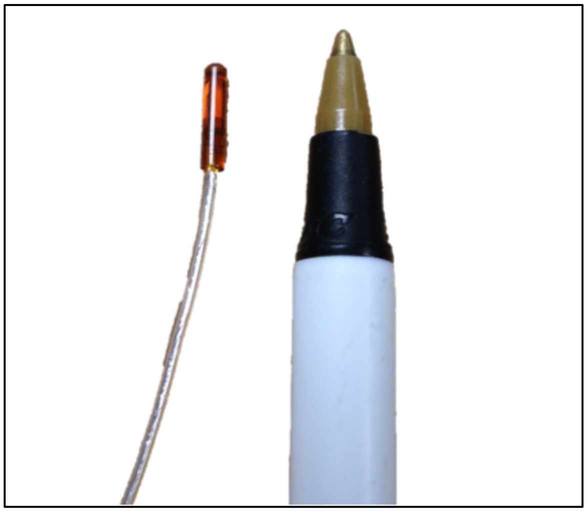

Figure 17. Ascension trakSTAR™ sensor size in relation to a ballpoint pen.

The battery-powered portable electric drill introduces a large amount of metal into the environment, making this the largest hurdle for the sensing components to overcome. The large amount of metal can distort the electromagnetic field emitted by the transmitter, resulting in inaccurate results in positioning. To diminish this issue, a plastic extension is attached to the drill, increasing the distance between the sensor and the metal, shown in Figure 18. This lessens the effect the metallic drill has on the sensor’s position output. A foam tube is wrapped around the artificial femur and secured with hook and loop straps for easy removal and access to the artificial femur. Figure 19 shows the femur wrapped in the foam tube, obstructing the direct view.

Figure 18. Simulator drill showing the extension for diminishing metallic interference in electromagnetic tracking.

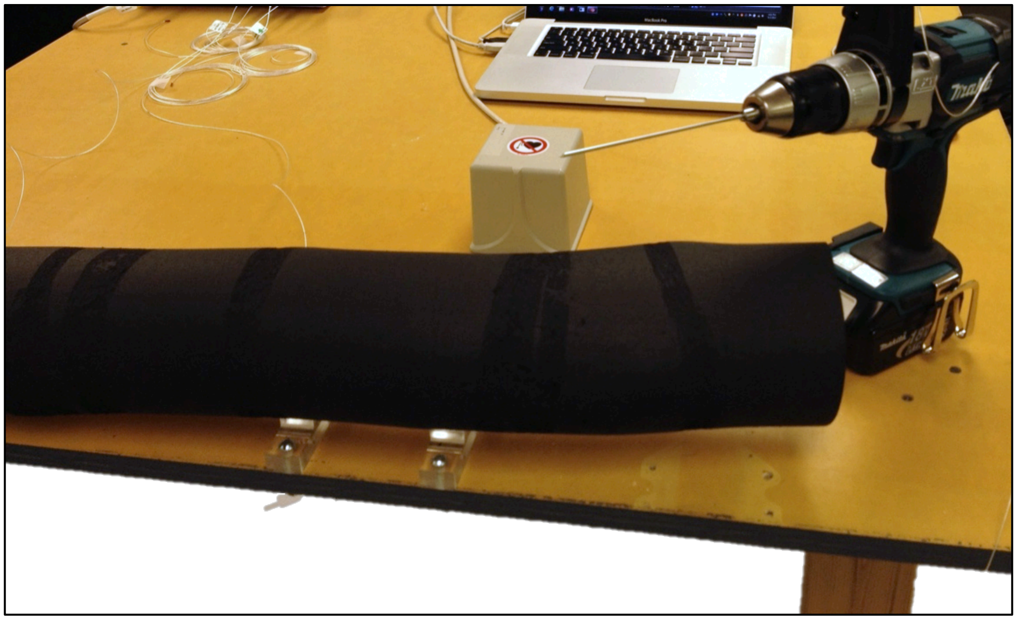

Figure 19. Artificial femur with foam tube for obstruction of direct view to trainee.

3.3.7 Software Overview

The simulator software consists of two major modules: sensor data collection and virtual fluoroscopy rendering. The sensor data collection occurs within the trakSTAR™ electromagnetic tracking device. Sensor data collection is enabled through modification of C language code written by the manufacturer of the electromagnetic tracking device, Ascension Technologies Corporation. The executable file for the electromagnetic tracking device provides data at the request of an external source through a data pipe. At request of the virtual fluoroscopy rendering software, the electromagnetic tracking software refreshes and updates the positional data of all attached sensors. The positional data is then sent to the fluoroscopy-rendering module. The primary advantage of this setup is enabling the tracking device to be swapped out for a different tracking device, whether optical motion-capture or a different electromagnetic tracking component. This allows the tracking system to be easily modified to suit the demand of the training regimen. Using the same logic, the software can be modified without disrupting the positional data input. This allows for the software to be updated or changed if desired, or to simulate different drilling procedures on other parts of the body. For example, without extensive reprogramming, the software could be modified to simulate drilling into a wrist or shoulder instead of a hip. A flowchart presenting the flow of data from mouse click to virtual fluoroscopy is shown in Figure 20.

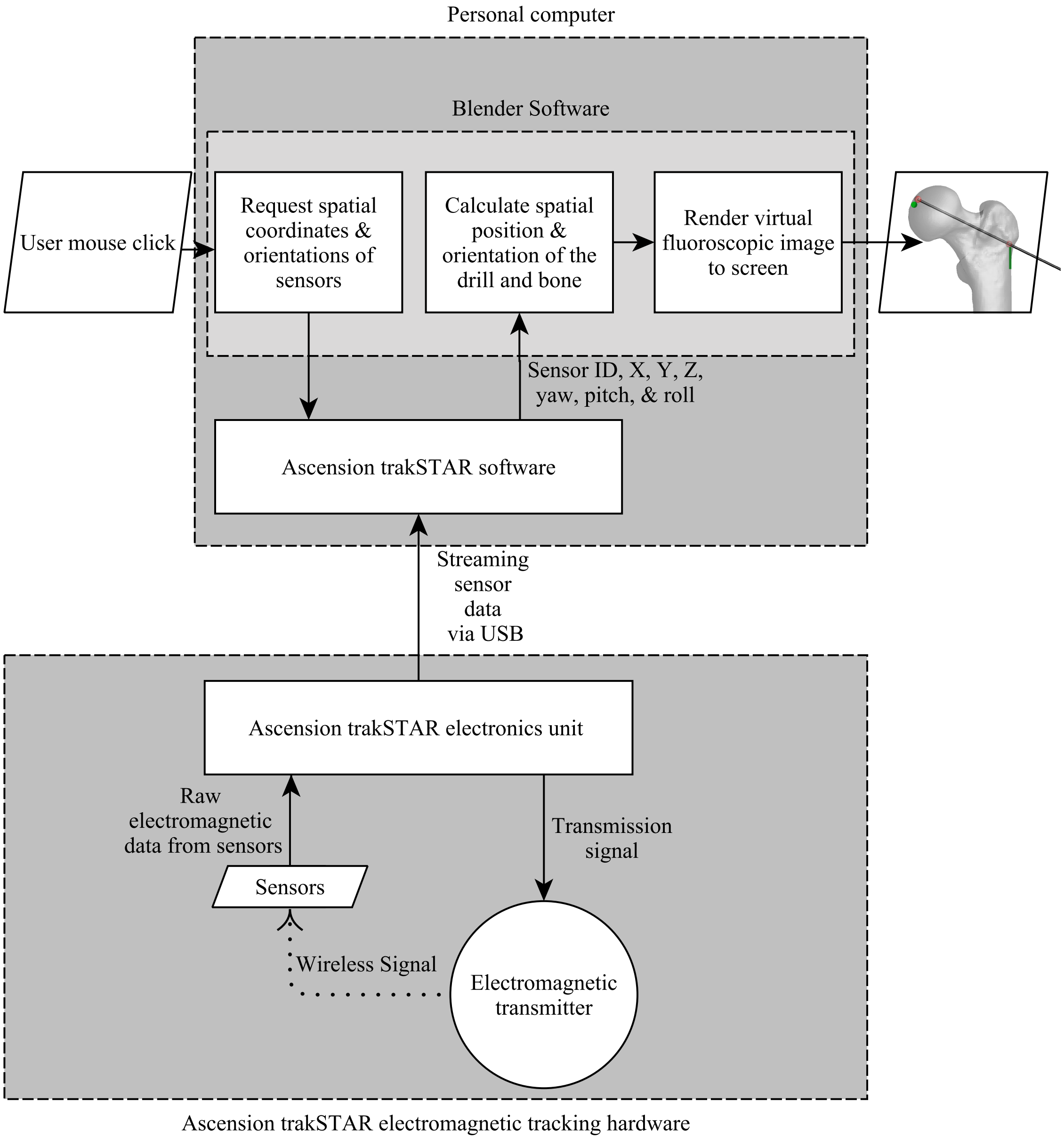

Figure 20. Flowchart of data modules for wire navigation simulator.

Generation of the virtual fluoroscopic images is accomplished through Blender (version 2.6), an open-source 3D software package. Blender is commonly used in animation films, 3D art, visual effects, and video game creation. The orthopaedic simulator utilizes the built-in game engine of Blender to expand the realm to simulation.

The game engine allows easy creation of an interactive environment with the user. In video games, interaction with the software is accomplished through keyboard and mouse input from the user. The simulator uses both mouse and keyboard input, but also uses input of positional data from the electromagnetic tracking device. Using inputs, the Blender game engine enables simple movement of 3D objects in the environment, including manipulation of lights and cameras in the virtual environment. For more advanced manipulation of the 3D environment, the game engine includes Python language scripting. Using Python, the user can change and manipulate 3D objects in code rather than using user interface buttons. Python scripts dominate the majority of the simulator structure. For example, keys on the keyboard and clicks on the mouse trigger are tied to specific Python scripts. Once the scripts are triggered, objects and properties are modified in the simulator environment.

The main component for manipulation in the virtual environment is the virtual femur. The virtual femur must accurately reflect the correct size and shape of a real human femur. In addition to size and shape, the virtual femur must be positioned correctly relative to drill position. This ensures that guide wire penetration of the artificial bone is accurately reflected in the virtual environment. For exactly that purpose, a laser scan was completed of the artificial femur used in the simulator. The digitization of the artificial femur, shown in Figure 21, serves as the 3D object for manipulation in the virtual environment.

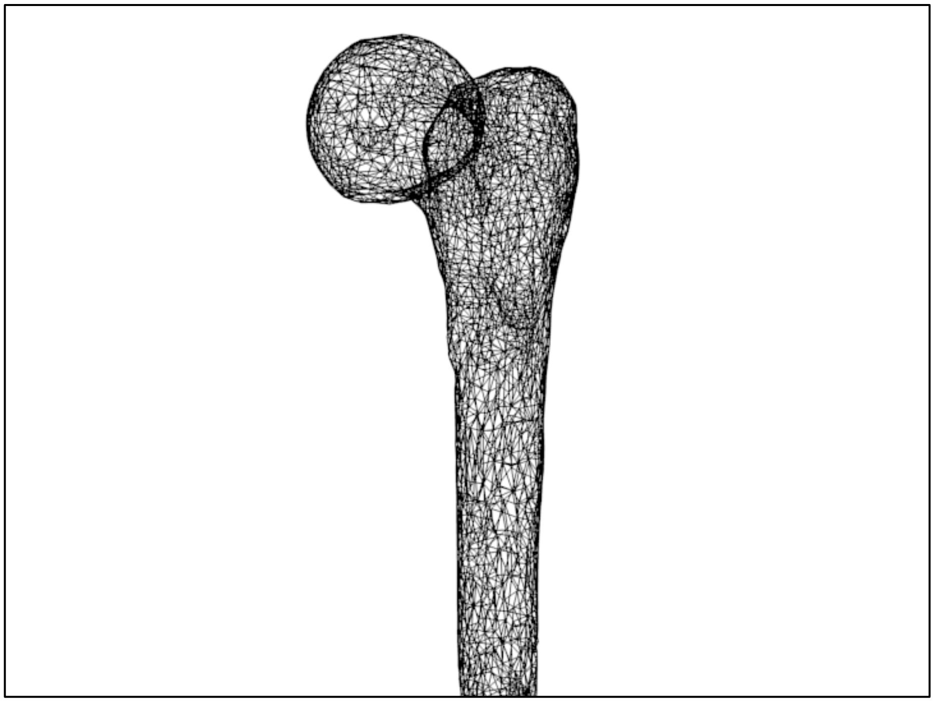

Figure 21. Three-dimensional laser scan of artificial femur.

3.4 Calibration of the Simulator

3.4.1 Introduction

One great challenge of creating the simulator is calibrating the electromagnetic tracking sensors inside the virtual environment. The original proposition of the simulator was to create a calibration procedure that could be used with different tools. In the case of this simulator, calibration procedures must be versatile enough to incorporate various sizes and models of drills. The calibration procedure is designed to be adaptable to future applications incorporating more procedures and tools. The simulator is also designed to be portable and compact. Allowing the simulator to be placed in various environments requires frequent recalibration due to different atmospheres and electrical interference. This created an additional need for the calibration procedure to be quick and versatile.

The calibration procedure employs a quick and robust method of correlating real world objects of simulator, including the drill and artificial bone, to their respective counterparts in the virtual environment. This entails a method of accurately placing the drill and artificial femur in correct locations relative to their respective sensors. Without mapping the 3D objects correctly in the virtual environment, the visual feedback given by the virtual fluoroscopy is an inaccurate representation of the drilling procedure in the real world. When done correctly, the virtual environment accurately matches up with the real world, as shown in Figure 22. The calibration consists of two main procedures. First, a procedure positions the drill and guide wire in an accurate position with respect to the corresponding electromagnetic sensor. The second step of calibration positions the bone correctly with respect to the drill and guide wire position.

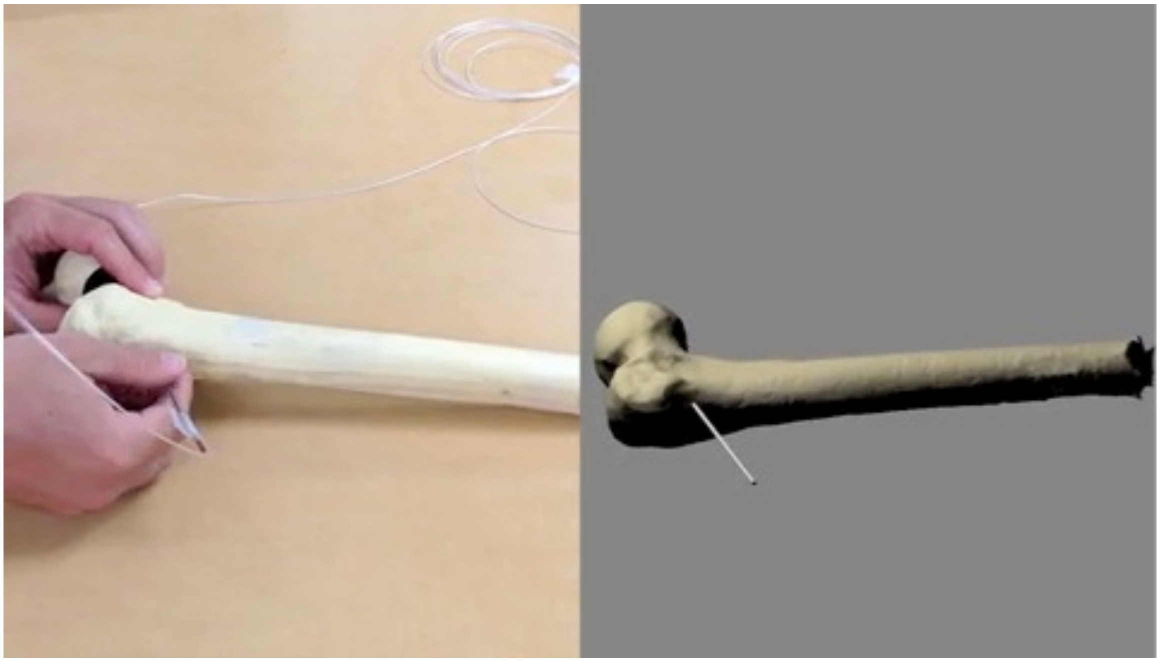

Figure 22. Side-by-side comparison of the correlation between real world and virtual world using calibration techniques.

3.4.2 Tool Offset

The vector from the electromagnetic sensor to the tip of guide wire is defined as the tool offset. This distance between the sensor and the guide wire can be manually measured with reasonable accuracy, but accurately calculating the vector through manual measurements is nearly impossible. The sensor's reference coordinate frame is embedded within the resin of the small sensor, therefore accurate orientation of the tool offset must be solved through inverse kinematics. The method uses collection of many data points and linear regression to find the best-fit vector from the sensor to the tip of the guide wire. The tool offset is used to determine the orientation of the guide pin with respect to the sensor. This provides the remaining tool offset data to accurately portray the guide wire in the virtual environment.

The procedure starts by placing the tip of the guide wire at a single point located on the simulator table with the electromagnetic sensor securely fastened to the drill. The drill is then rotated and manipulated while keeping the guide wire tip fixed at a specific location (see Figure 23). This offline procedure need only be repeated when a different length guide wire is placed in the drill, or the sensor position on the drill is changed. During this procedure, a Python script collects positional data from each of the sensors and dumps coordinates and orientations into a data file. The software allows the user to specify the number of data coordinates used in the calibration procedure. The collection of more points minimizes slight errors in positional data, but also increases the calibration time. Between 400-1500 positional data points, depending on the metallic interference of the environment, yields an accurate tool offset.

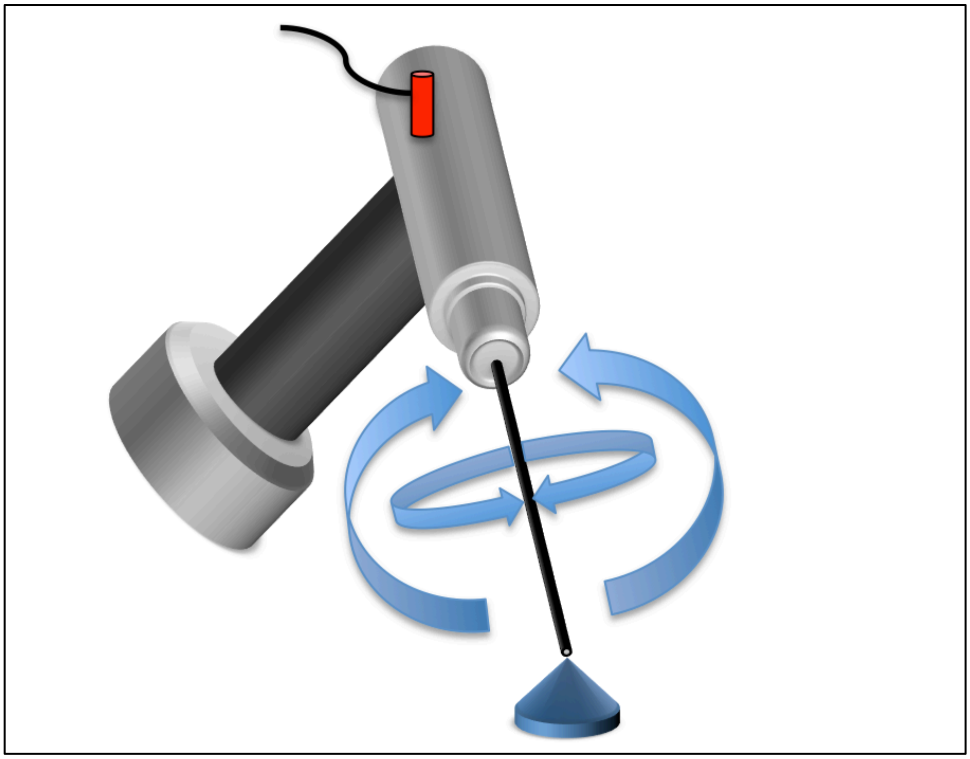

Figure 23. Graphical representation of tool offset calibration technique.

Inverse kinematic equations use the collected data points to calculate the tool offset. The sensor collects many data points about the stationary tip. The calibration solves the best-fit solution of the center of all the collected points. A visualization of the calibration procedure is shown in Figure 24.

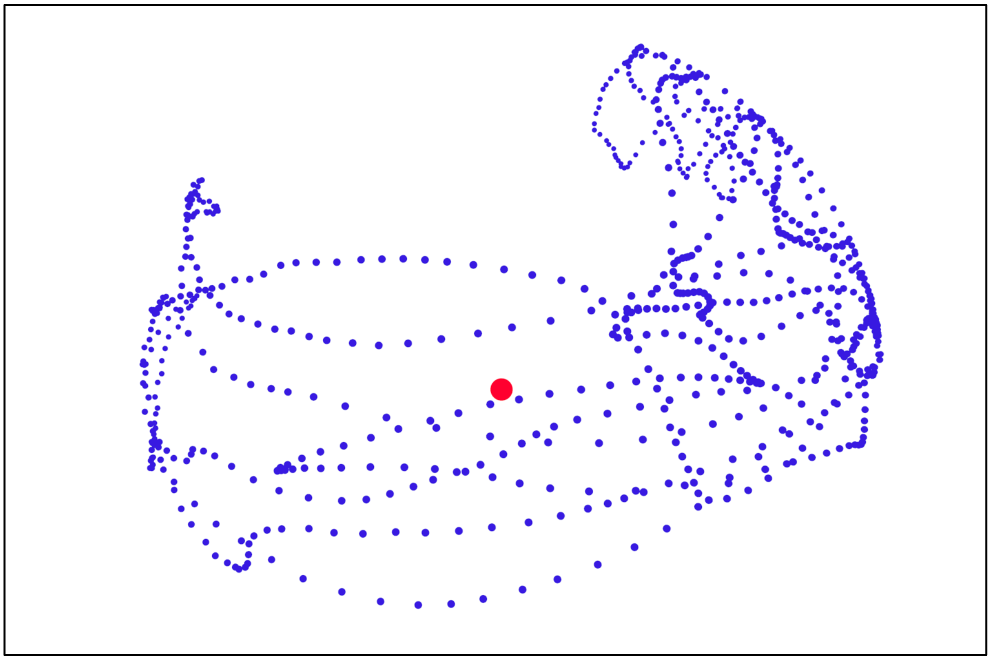

Figure 24. Visual representation of the tool offset calibration procedure. Data points are collected as the drill is manipulated, shown in blue. These data points are then regressed to determine the best-fitting position of the stationary point and the corresponding best-fitting tool offset.

The unknown vector from the tip of the guide wire to the transmitter remains constant due to the constraints of the calibration technique. The vector from the transmitter to sensor attached to the drill is known, for this is the positional data collected during the tool offset calibration procedure. The unknown vector from the sensor to the guide wire tip is the vector of interest for the tool offset. The kinematic representation of the vectors is shown in Figure 25.

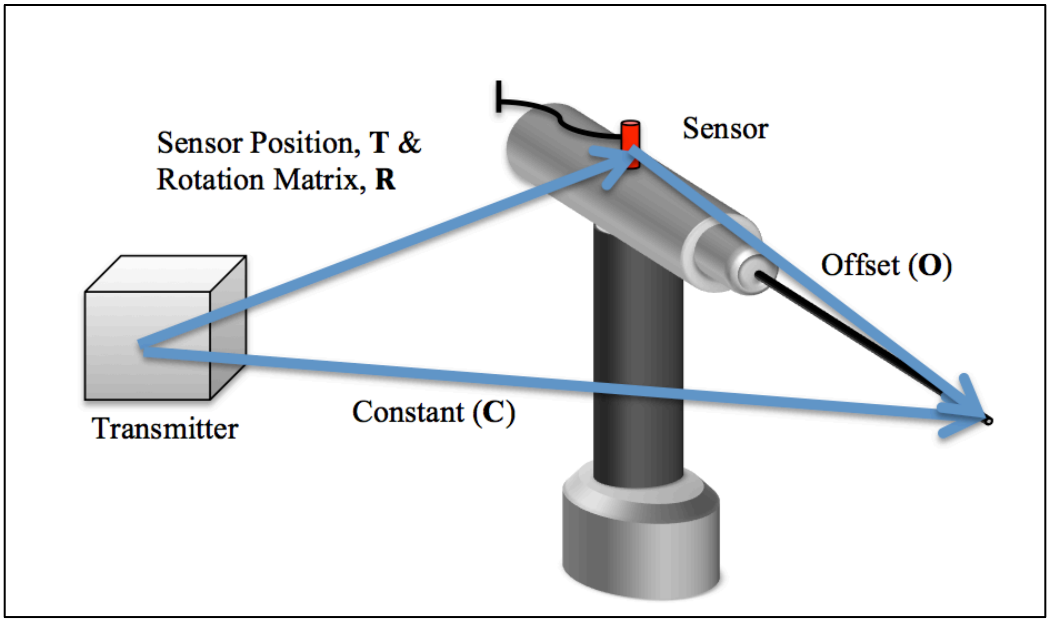

Figure 25. Graphical representation of tool offset calibration calculation.

Using Figure 25, a vector equation is established and shown in equation (2). **R** is the Euler rotation matrix of the sensor orientation. The Euler rotation matrix is calculated from the orientation data provided by the sensor data collected during calibration. **O** is the unknown offset vector from the sensor to the tip of the guide wire. **T** represents a single positional vector collected during the calibration procedure. Lastly, **C** is the vector from the transmitter to the tip of the guide wire, which remains constant during the calibration procedure.

$$\mathbf{RO} + \mathbf{T} = \mathbf{C} \qquad (2)$$

Bringing all the unknowns to one side of the equation rearranges equation (2) to become equation (3).

$$\mathbf{RO} - \mathbf{C} = -\mathbf{T} \qquad (3)$$

Expanding the vectors and Euler rotation matrix expand and combining the unknowns into a single vector results in equation (4).

$$\begin{bmatrix} R_{11} & R_{12} & R_{13} & -1 & 0 & 0 \\ R_{21} & R_{22} & R_{23} & 0 & -1 & 0 \\ R_{31} & R_{32} & R_{33} & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} O_x \\ O_y \\ O_z \\ C_x \\ C_y \\ C_z \end{bmatrix} = \begin{bmatrix} -T_x \\ -T_y \\ -T_z \end{bmatrix} \tag{4}$$

Equation (4) represents the six unknowns and three equations associated with each data point collected during the calibration procedure. This equation takes the general form of the simplistic linear equation **Ax = b**. These equations grow by a factor of three for every additional data point collected during tool offset calibration, while simultaneously, the number of unknowns remains constant. By collecting n-positional data points, the total number equations amount to 3n, but the number of unknowns remains at six. Combining multiple positional data points, equation (5) represents equation (4) including n-data points. Due to the structure of the equation, linear regression is performed to find the best-fit tool offset vector for the positional data collected. Solution to the linear regression results is found by equation (6).

$$\begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \\ \vdots \\ \mathbf{A_n} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \\ \vdots \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b_1} \\ \mathbf{b_2} \\ \vdots \\ \mathbf{b_n} \end{bmatrix} \tag{5}$$

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \tag{6}$$

Upon solving the equation, a text file is written to the Blender directory. This text file contains vector information that is used by the virtual fluoroscopy software. The vector from the sensor to the tip of the guide wire is calculated, but the orientation of the guide wire still needs calibration, leading to the second step of the tool-offset calibration. The remaining calibration is done inside the Blender game engine environment.

3.4.3 Orienting the Guide Wire

The preceding calibration calculates the vector from the sensor to the tip of the guide wire, but the orientation of the guide wire in respect to the sensor remains unknown. Determination of the orientation of the guide wire is accomplished with the use of a second electromagnetic sensor. The second sensor is placed on the guide wire with its longitudinal axis pointed directly inline with the guide wire's longitudinal axis, shown in Figure 26. At this point the simulator registers the orientation of the secondary sensor and maps the guide wire rendered in the 3D environment to the respective orientation.

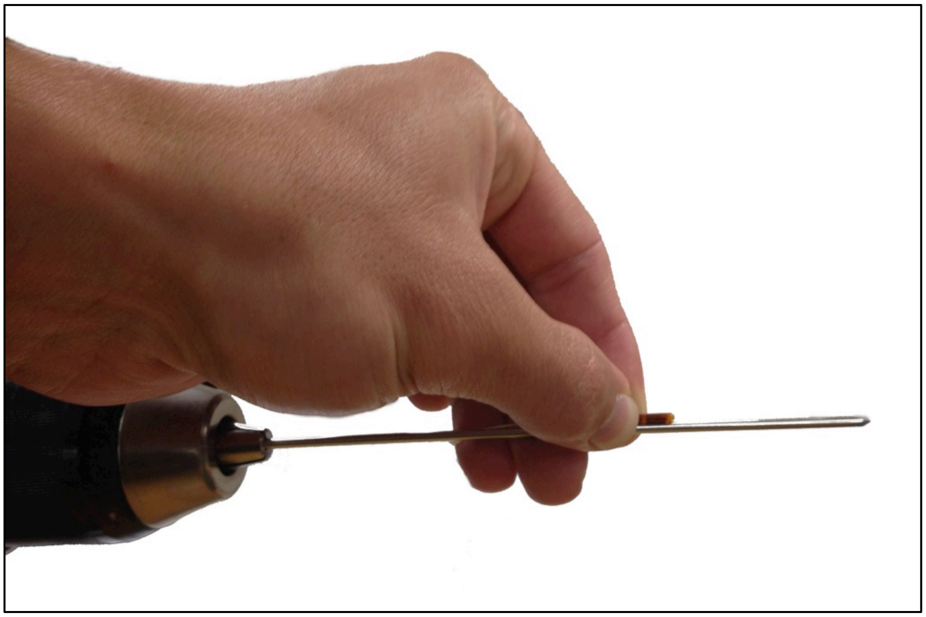

Figure 26. Procedure for orienting the guide wire in the virtual environment.

3.4.4 Artificial Femur Registration

Upon orienting and calculating the offset in the virtual environment, the drill component of the simulator is completely calibrated for simulator implementation. Using this calibrated instrument, the artificial bone is accurately mapped to the virtual environment through an additional procedure. This procedure insures that haptic feedback to the user accurately reflects collisions in the virtual environment. The electromagnetic tracking system has small inconsistencies, noticeably accumulating over large positional changes. The proximal femur is the reference position for the guide wire; therefore, it must be the most accurate region for haptic feedback purposes. To combat dimensional inaccuracy issues, the bone registration procedure aims to calibrate the proximal femur region rather than averaging the accuracy over the entire femur. This ensures the drilling region is the most accurately calibrated area.

The three calibration points used for a touch-off registration procedure are shown in Figure 27. The first registration point is located at the center of lateral cortex of the proximal femur. The tip of the guide wire touches this point and the software records the positional data. The positional data is used to translate the 3D model of the artificial femur inside the virtual environment to the correct corresponding location. Second, the apex of the artificial femur is touched with the guide wire, representing the second registration point. The software records the positional data. A vector is constructed from the previous position touched on the lateral cortex to the apex of the femur correctly in the virtual environment. This corresponding vector is mapped from the real world to the virtual environment. The final and third registration point in the calibration procedure is the apex of the lesser trochanter on the artificial femur. This final point rotates the femur about the previously constructed vector in the virtual environment to accurately orient the virtual femur. Completing this calibration correctly maps the virtual bone with respect to the sensor on the drill. The artificial femur is held stationary throughout the registration procedure, but after completion of the procedure, the femur can be dynamically repositioned without losing calibration. The second sensor attached to the artificial femur allows the capability, although not yet utilized, for the user to rotate and move the femur throughout the procedure.

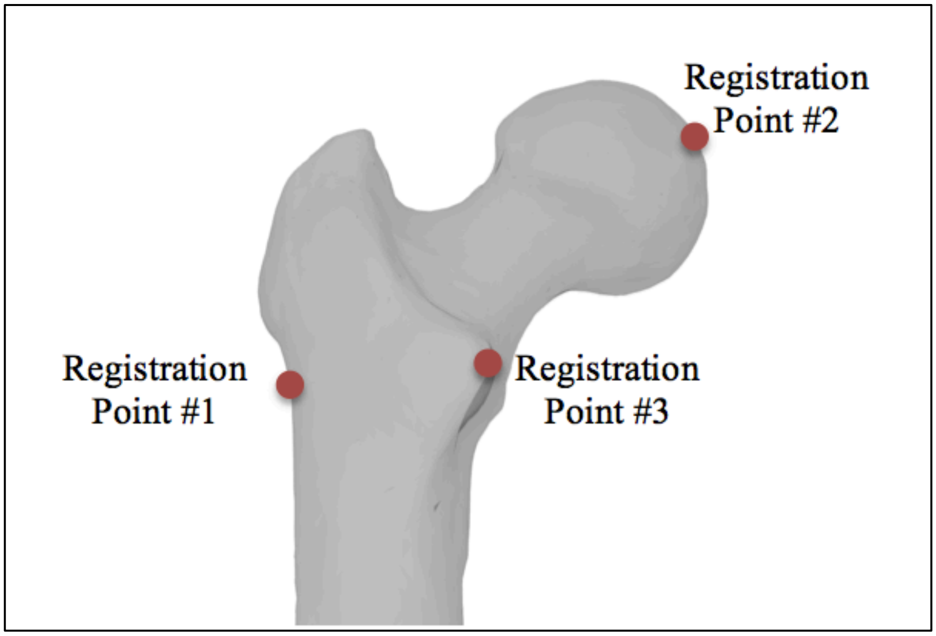

Figure 27. Diagram of three-point registration calibration procedure.

Commonly, the second sensor is not directly attached to the bone, it is often attached to the simulator table. This is possible as long as there is no relative motion between the sensor and the bone. The advantage of this setup technique allows quick swapping of artificial bones in the simulator without the need to detach the sensor from the previously drilled bone and reattached to the new artificial femur for the next trial. Attaching the second sensor to a table can also be used to mitigate recalibrating the position of the artificial femur when swapping to a new femur. If the new femur is placed in the exact same position of the previous femur there is no need for recalibration. This is accomplished with an external indexing device for precise repeatable placement of the artificial femur.

3.4.5 Accuracy of the Tracking Sensors

Simulator precision is important to assess, and is a component contributing to ensuring that practice on the simulator develops psychomotor skills closely related to the task. It also documents system performance for other researchers. The principal limitation of the simulator is related to the precision of the wire tip position tracking. The precision is a function of the resolution of electromagnetic tracking sensors, as well as fitting errors due to algorithm calibration. Quantification of guide wire tip location precision after calibration documented the overall reliability of the tracking system. The tip of the guide wire was fixed at a single point, and the drill was moved to 800 different orientations around the tip of the guide wire. The collected coordinates are shown graphically in Figure 28 (deviations from the origin are errors in measurement). This procedure was used to assess the cumulative calibration errors and electromagnetic tracking sensor resolution. The average error in the x, y, and z-directions were 1.50 millimeters, 2.29 millimeters, and 1.55 millimeters, respectively. The average overall spatial error was 3.57 millimeters. The maximum deviation of the 800 points collected was 9.45 millimeters. A calculated 95% confidence interval for the net spatial errors extends from 3.44 millimeters to 3.70 millimeters. The absolute spatial error of human proprioception ranges between one and two centimeters [151, 152]. This data demonstrates spatial errors range from 3.4 millimeters to 3.7 millimeters due to calibration and tracking inaccuracies. The simulator accuracy exceeds the accuracy of the human senses while demonstrating consistency. This study shows that the simulator is accurate enough for training purposes.

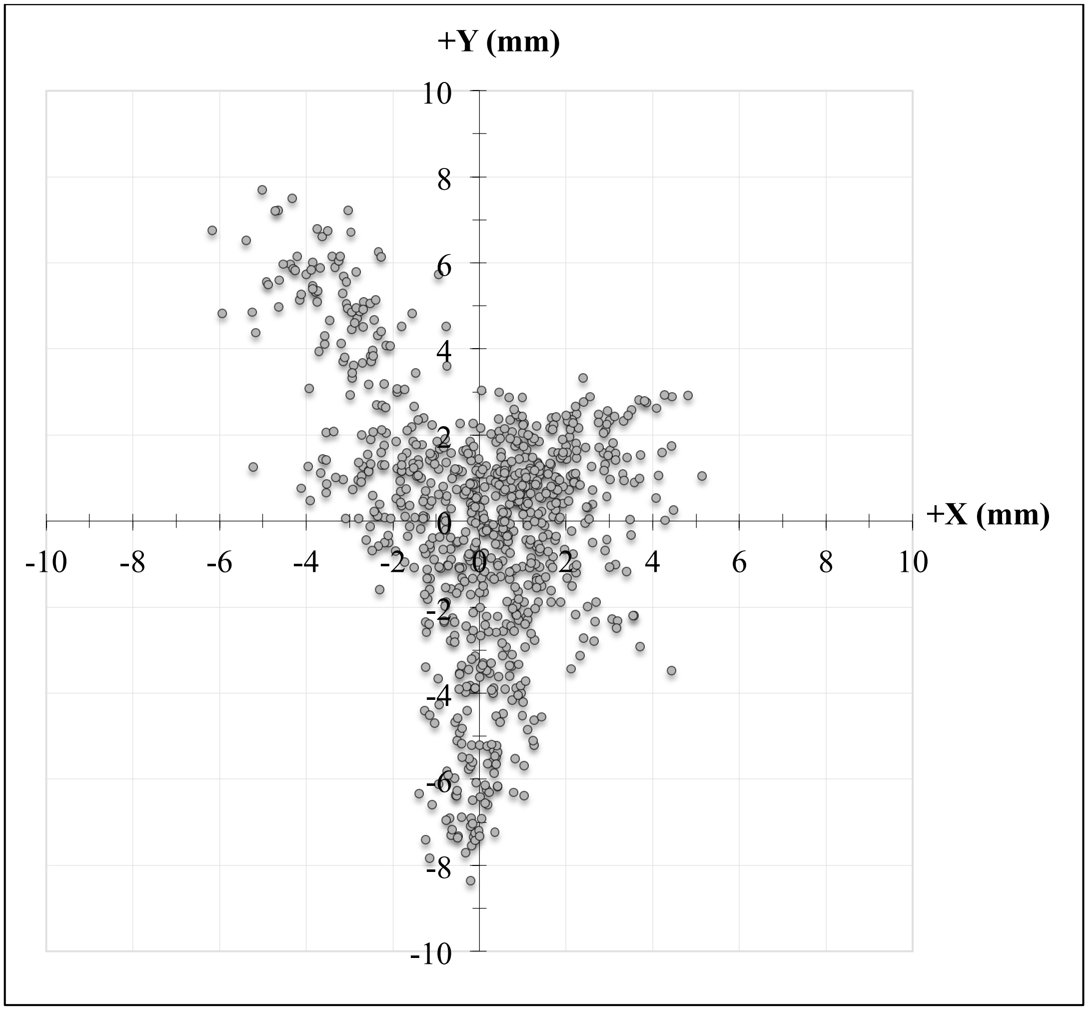

Figure 28. Eight hundred 3D coordinates for assessing the accuracy of the simulator (z-direction omitted in graphic).

3.5 Operating the Simulator

3.5.1 Introduction

Simplicity is the aim of the simulator for operation. Once calibrated, the simulator requires only mouse clicks and a few keyboard presses to run. For implementation within a training curriculum, it is important that first time users can operate the simulator. Although normal operation is simple, the simulator also provides advanced features and expandability for advanced users through multi-button keystrokes on the keyboard. The current setup of the simulator is set for two participants, one completing the orthopaedic drilling, while the other runs the virtual fluoroscopy, acting like a radiology technician in the operating room. Starting the simulator is done by simply pressing “P” (for “Play”) on the keyboard. The Esc key exits the simulator trial.

3.5.2 Virtual Fluoroscopy Operation

The main function of the software is to generate virtual fluoroscopy, mitigating the need for radiology tools in training. The simulator operates by simulating static fluoroscopic images. This simulator gathers information from the sensors at an instant in time and constructs the respective virtual fluoroscopic image. This static image remains on the screen in a paused state until the next image is rendered. The simulator is generally paused, briefly resuming for new image generation. The simulator can present either the anteroposterior (AP) or the lateral fluoroscopic view (see Figure 29). The two static views are available through the left and right mouse buttons. Pressing the left mouse button generates the AP view in virtual fluoroscopy, while pressing the right mouse button generates the lateral view. Pressing the middle mouse button ends the simulation trial.

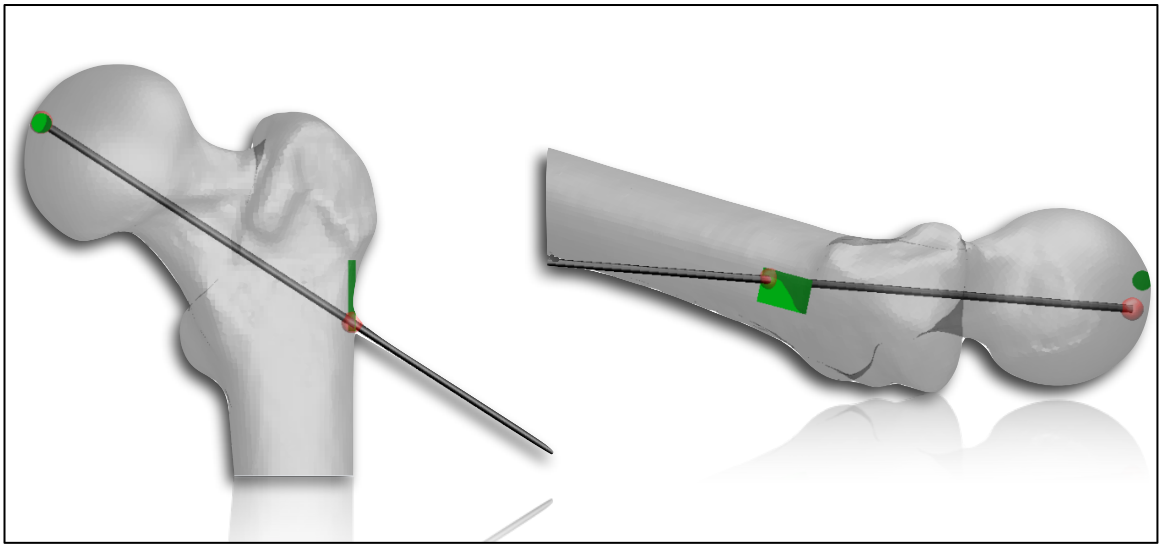

Figure 29. Sample virtual fluoroscopy with AP (left) and lateral (right) views.

3.5.3 Simulator Customization and Options

Several advanced features are incorporated in the simulator. The normal operation uses static virtual fluoroscopic images due to the static nature of actual fluoroscopy, but a dynamic mode is available. This dynamic mode continuously streams positional data to the game engine software from the electromagnetic sensors. This mode is helpful when introducing a trainee to the simulator for the first time. Using the live video stream, the trainee easily associates drill movements and their corresponding actions within the virtual fluoroscopy. It is important to note that activating the drill generates electromagnetic noise disrupting positional accuracy. This dynamic mode is also used during calibration. Most of the other keyboard commands can only be used in the calibration mode. This calibration mode allows the technician to setup and calibrate the simulator while seeing the dynamic movement of the virtual environment. This provides the ability to recognize if the simulator is calibrated properly.

Different surgeons prefer slightly different lateral views of the hip and variations in patient and fluoroscope setup cause slight view differences in the lateral views available during surgery. Due to the variability in lateral views, advanced options are added to the simulator giving the user freedom to manipulate the angle of the lateral view. The software allows setting and saving four personalized lateral views prior to simulation. This accommodates the advanced user, who may use slightly different lateral views to ensure accuracy of the guide wire. Figure 30 shows samples of saved options for the lateral views in the simulator. Although a neutral angle view is the default setup, personalizing the lateral view is available.

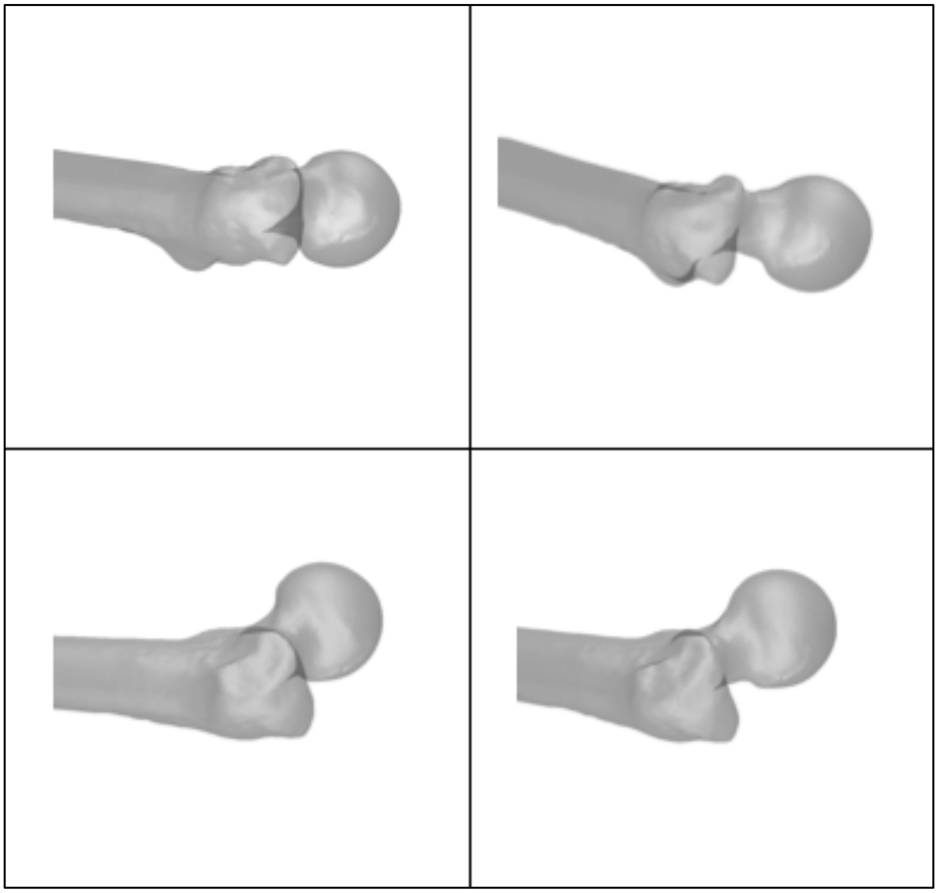

Figure 30. Four lateral view options of virtually generated fluoroscopic images.

In addition to adjustment of the lateral views, advanced users can fine-tune the tool offset. Using keystrokes on the keyboard, the angle of the guide wire may be adjusted and manipulated. This feature allows the tool offset of the guide wire to be manually adjusted when needed. This is helpful when electromagnetic interference is present during calibration. It is also a valuable feature when tool length is changed.

The advance features of the simulator add benefits to the setup, introduction, and operation of the simulator. The simulator is intended to be simple and intuitive, but the advanced features are added to accommodate and troubleshoot a variety of situations. These features are available through keystrokes on the simulator keyboard, but rarely need to be evoked by the average user. A layout of the simple keyboard commands is shown in Figure 31.

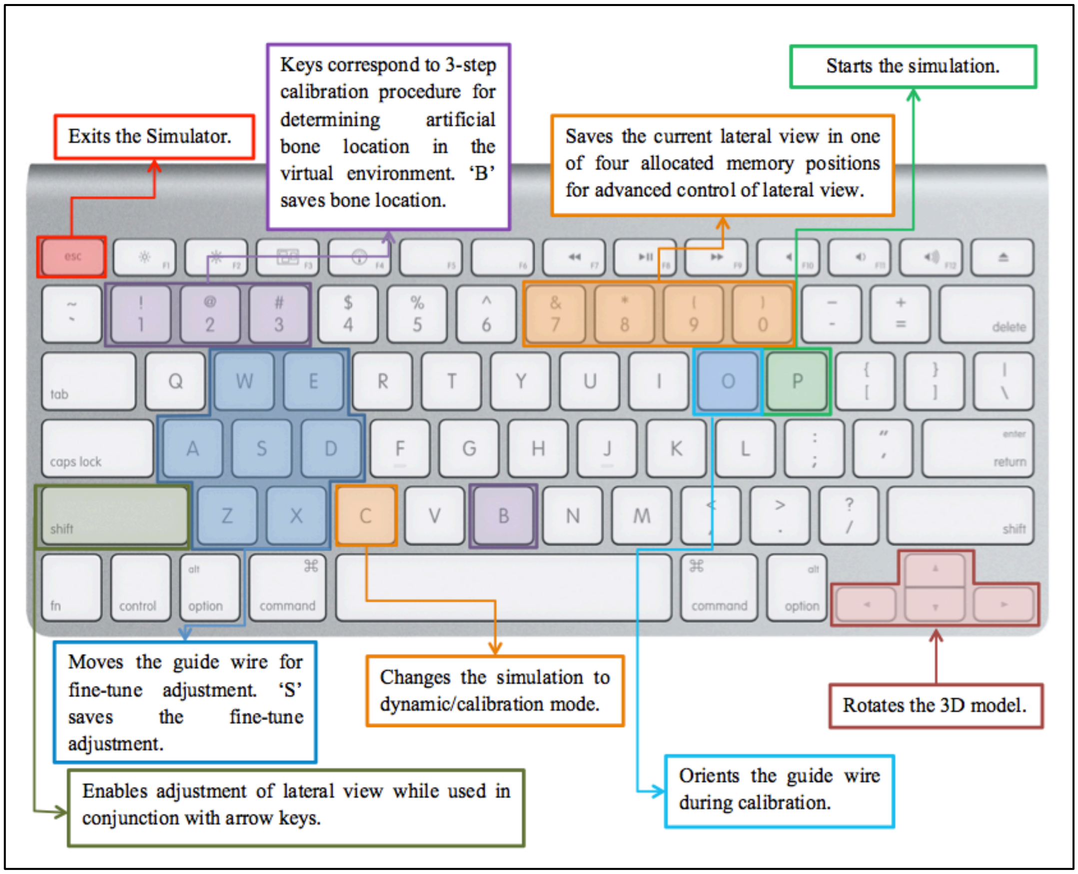

Figure 31. Keyboard command layout for simulator.

3.6 Data Collection and Feedback

3.6.1 Data Collection

The ability to collect data throughout the simulator trial is a main advantage over other forms of practice, including actual surgery in the operating room. The simulator enables data collection for assessment and feedback of the trainees on command. At the start of a simulator trial, a unique file directory is created in the simulator file structure. Simultaneously, as the directory is created, a common-separated value (.csv) spreadsheet file is created in the directory for trial data storage. As the trial progresses, each virtual fluoroscopic image is also saved to the unique trial directory with a filename identifying the view used and the exact time of the creation of the image. Each image is recorded in a spreadsheet file and is time-stamped. In addition to the time of each image, the tip-apex distance is recorded along with the coordinates of the entry point of the guide wire. When the simulator trial finishes, two final images are recorded in the directory and the finishing time is written to the spreadsheet.

3.6.2 Feedback

At completion of a trial, the trainee verbalizes his or her completion. At this time, the virtual guide wire is finalized in the virtual femur and is untethered from the positional control of the electromagnetic sensor. The guide wire is removed from the artificial bone and the trainee views the result on the computer screen. Upon completion, the virtual fluoroscopy becomes a rotatable 3D virtual model of the proximal femur. Several highlights on the virtual model become present when the trial ends. First, the tip of the guide wire and entry location of the guide wire in the femur is highlighted in red. An area of acceptable entry area determined by expert surgeons is highlighted in green. Finally, the apex of the femur is highlighted in green so that the trainee comprehends the tip-apex distance visually. The arrow keys on the keyboard allow the user to manipulate the virtual femur with the resulting guide wire embedded in the femur. This rotatable model supplements the AP and lateral images given throughout the trial for full comprehension of the three-dimensional position of the guide wire. Figure 32 shows an interpretation of available three-dimensional feedback.

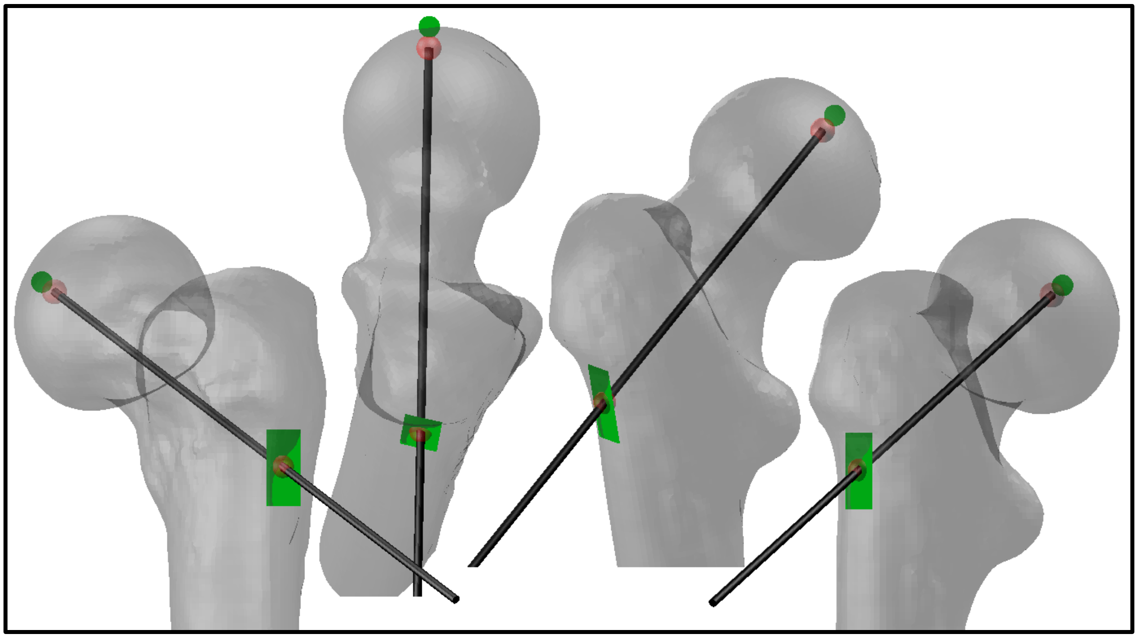

Figure 32. Three-dimensional capability for training feedback.

After viewing the 3D model, the trainee has easy access to the number of fluoroscopic images used, the total time of the trial, and accuracy of the guide wire placed in the proximal femur. In addition to the trial result data, the trainee has access to review the trial image-by-image along with the accuracy progression throughout the trial. A sample of the data collected throughout a trial is shown in Figure 33.

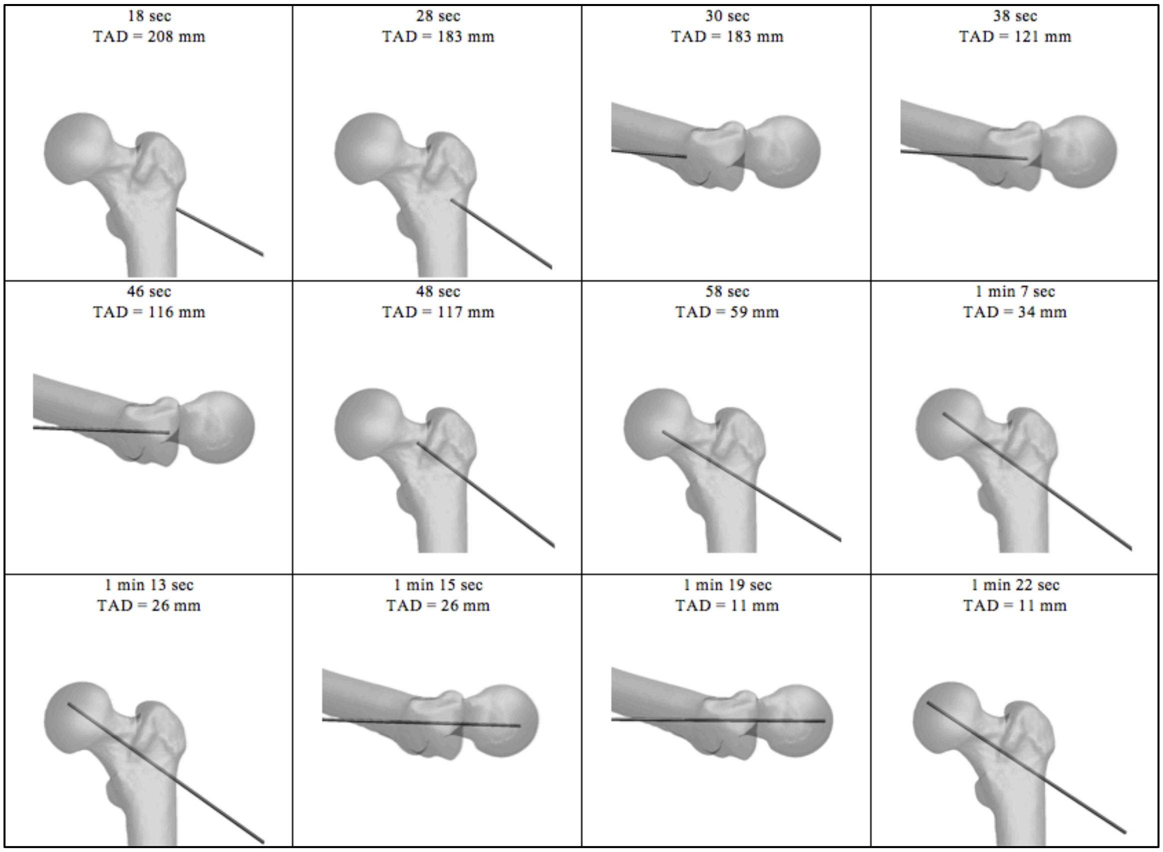

Figure 33. Data sample collected throughout a single trial.

3.7 Discussion

This chapter provides evidence that an effective orthopaedic simulator can be developed using augmented reality. The development demonstrates an alternative method for simulator creation by using augmented reality rather than the more frequently used haptic feedback devices. The design emphasizes stimulus-response cues rather than the pursuit for complete realism. In addition, this research gives future simulator developers a detailed account of calibration techniques for future augmented reality simulators. Overall, this section advance knowledge in simulator creation by implementing a novel application of technology in creating a simulator. This simulator hopes to serve as a device that is further advanced and improved by future developers.

CHAPTER 4 – AIM TWO: CONSTRUCT VALIDITY OF THE AR WIRE NAVIGATION SIMULATOR

4.1 Introduction

The second aim of this research is to demonstrate construct validity of the developed simulator. This section details experimental simulator trials performed by engineering undergraduate students, both novice and experienced orthopaedic surgical residents, and faculty orthopaedic surgeons. This chapter intends to detect quantitative differences in performance between surgical skill levels. Statistically discriminating evidence indicates that the simulator objectively measures the surgical skill it claims to measure. This establishes the construct validity of the simulator. The evidence strongly advocates for use of the simulator as a valuable training and assessment tool in wire navigation of the proximal femur.

4.2 Hypothesis

It is hypothesized that the experienced surgical group will outperform novices in wire placement accuracy, time to complete wire navigation, number of images used throughout the trial, and number of drilling attempts necessary to place wire. These results will demonstrate construct validity for the developed augmented reality wire navigation simulator.

4.3 Methods

4.3.1 Participants

Forty subjects participated in the study. The subjects involved of three groups: 22 undergraduate engineering students, nine first-year orthopaedic residents, and nine experienced surgeons. The experienced group consisted of five fourth-year orthopaedic residents and four faculty surgeons. The University of Iowa Institutional Review Board (IRB) approved the study. All subjects were given basic instructions regarding wire navigation and were provided with simulator demonstration prior to beginning the experiment. Participants then completed three consecutive trials on the augmented reality wire navigation simulator, receiving feedback between the individual trials. An example of one trial is shown in Figure 34.

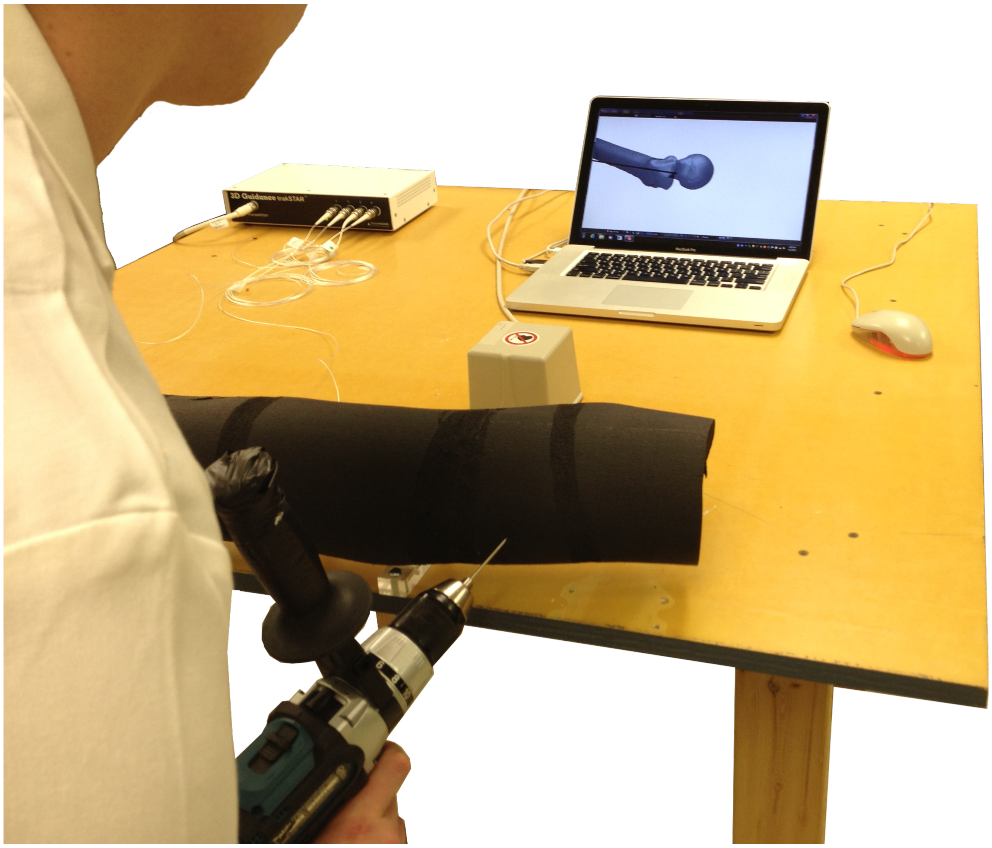

Figure 34. Example of a trial on the augmented reality wire navigation simulator.

4.3.2 Performance Assessment

The TAD, guide wire entry location, trial duration, number of images, and number of drilling attempts were all utilized to assess performance. The simulator recorded all dependent variables, except for the number of drilling attempts. The number of drilling attempts was calculated from the images and corresponding TAD data collected during the trial. To qualify as an additional drilling attempt, two pieces of criteria were required. First, an image showed an increase of 15 millimeters or more in the TAD from the previous image. Second, the wire also had to previously achieve penetration of the femur's lateral cortex.

Performance differences were compared between subject categories using a repeated measures multivariate analysis of variance (MANOVA). A statistically significant MANOVA was followed by a series of one-way ANOVAs across all dependent variables. Post-hoc t-tests were completed to compare dependent variable means for statistically significant one-way ANOVAs. Performance differences were interpreted as evidence in support of the simulator's construct validity for the wire navigation task.

4.4 Results

4.4.1 Construct Validity

Forty subjects were used in the statistical analysis. A repeated measures MANOVA revealed significant multivariate main effects in test subject experience levels, revealing Wilks' Lambda = 0.558, F (10, 70) = 2.375, p = 0.017, partial eta squared = 0.253. Power to detect the experience effect was 0.907, confirming that dependent variables can detect quantifiable differences between undergraduate students, novices, and experienced surgeons on the simulator. This confirms the hypothesis of the simulator in terms of demonstrating construct validity. Due to the significance of the overall test, the univariate main effects were also examined. A full list of univariate tests is shown in Table 3.

Significant univariate main effects for experience were obtained for TAD, with F(2, 39) = 3.721, p = 0.033, partial eta squared = 0.160, power = 0.648; and the number of drilling attempts, with F(2, 39) = 5.771, p < 0.01, partial eta squared = 0.228 , power = 0.841. Univariate effects for TAD and drilling attempts are shown in Figure 35 and

Figure 36, respectfully. Due to these significant univariate effects, further post-hoc mean comparisons were completed. Significant pairwise differences were obtained in the TAD between the experienced surgeons, first-year orthopaedic residents, and undergraduate engineering students. The mean for TAD in the experienced surgeons was 11.9 millimeters, while the means for the first-year residents and engineering students were 21.9 millimeters and 23.1 millimeters, respectively. This yielded a p-value of 0.011 between experienced surgeons and engineering students. The TAD also yielded a p-value of 0.043 between experienced surgeons and first-year orthopaedic residents. In addition to TAD, a significant means variation was found between the number of drilling attempts made by engineering students and first-year orthopaedic residents ($p=0.002$). First-year orthopaedic residents averaged 1.076 drilling attempts and first-year orthopaedic residents averaged 2.333 drilling attempts. These statistics demonstrate how experienced surgeons more accurately drill a guide wire and require less time to perform the task on the simulator than novices. These results demonstrate preliminary construct validity of the wire navigation task on the simulator.

Table 3. Univariate effects between experience groups (* indicates a significant effect, $p<0.05$).

Measure	F	Significance (p)	Partial η^2	Observed Power
TAD*	3.721	0.033	0.160	0.648
Entry Location	0.058	0.943	0.003	0.058
Images	2.304	0.113	0.106	0.440
Duration	2.910	0.066	0.130	0.536
Drilling Attempts*	5.771	0.006	0.228	0.841

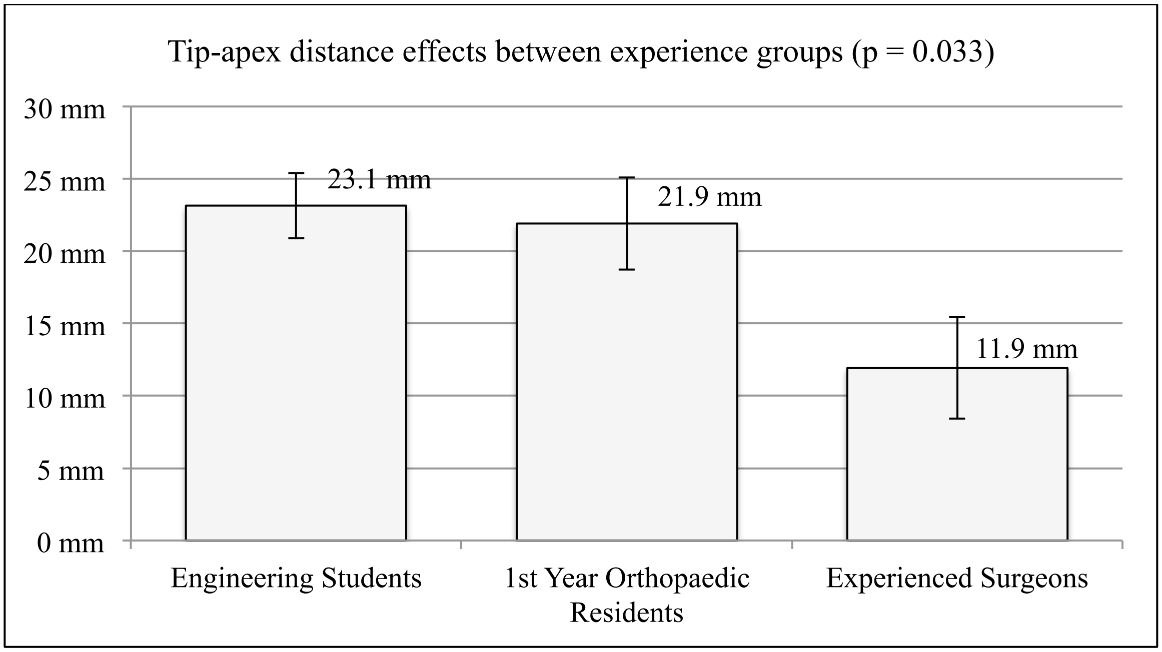

Figure 35. Significant univariate effects in TAD between subject groups.

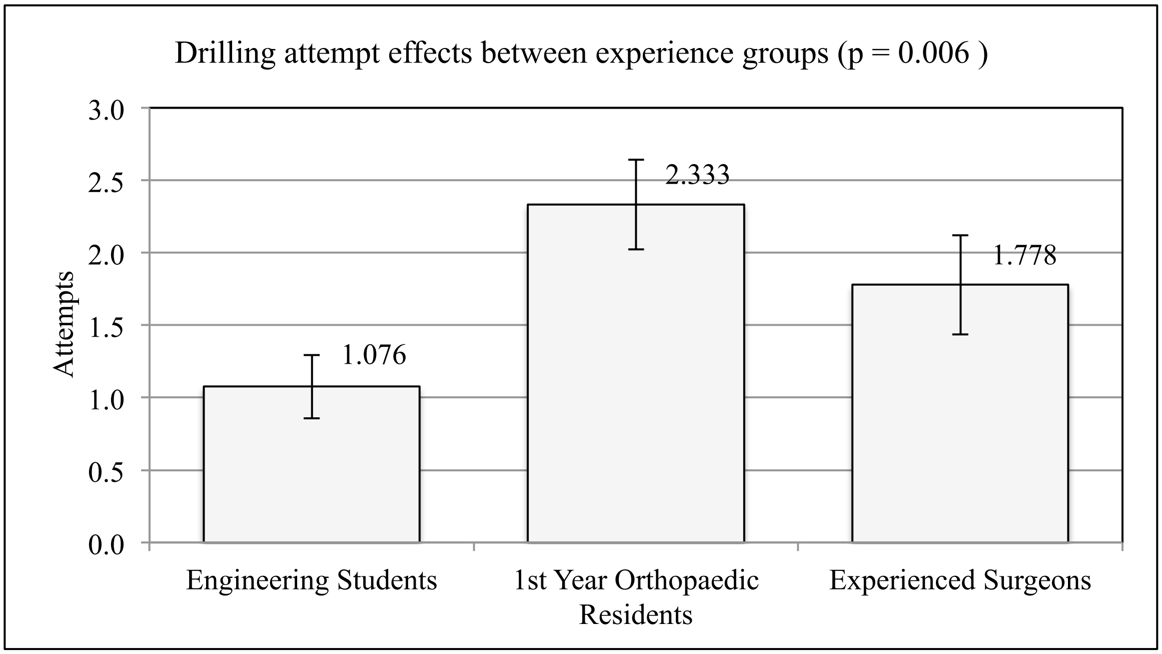

Figure 36. Significant univariate effects for drilling attempts between engineering students and first-year orthopaedic residents.

4.4.2 Additional Findings

In addition to the hypothesized results, other noteworthy findings appeared in the experimental trials. This experiment utilized a between-subject experimental design, controlling the experience of participants, yet there were also interesting within-subject effects. A repeated measures MANOVA revealed significant multivariate main effects for the trial number, revealing Wilks' Lambda = 0.443, F(10, 30) = 3.922, p = 0.002, partial eta squared = 0.567. Power to detect the trial effect was 0.981. Significant univariate main effects for trial number were obtained for the duration of the trial, with F(1.904, 74.245) = 10.405, $p < 0.001$, partial eta squared = 0.211, power = 0.982. This result shows that subjects were improving with each subsequent trial. Figure 37 shows the duration of each subject group over the three trials. In addition to duration, the TAD was approaching significant univariate main effects for trial number, with F(1.68, 65.532) = 3.721, p = 0.053, partial eta squared = 0.072, power = 0.522.

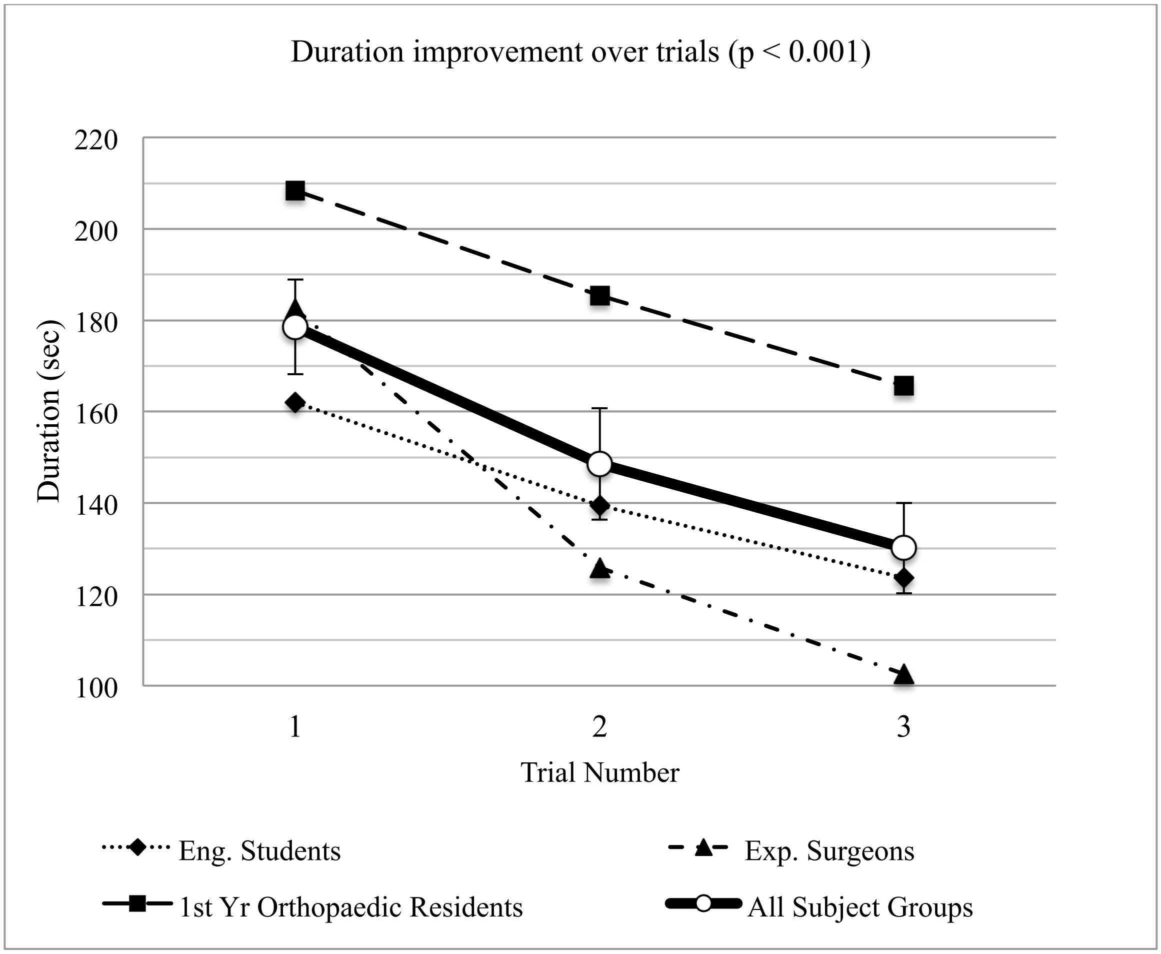

Figure 37. Improvement over time in three trials of the orthopaedic wire navigation simulator.

4.5 Discussion

Wire navigation is a skill-intensive task critical to the success of hip fracture surgery and to the longevity of the repair. Other researchers agree that the orthopaedic field shows promise for improving guide wire placement through simulation [102]. Validity studies will help aid the implementation of surgical simulators into orthopaedic training programs.

The goal of the study was to assess and validate the AR wire navigation simulator. Unlike previously developed wire navigation simulators using haptic feedback devices, this study used realistic, real-world objects and electromagnetic tracking technology to construct a novel AR wire navigation simulator. In addition, the study endeavors to demonstrate construct validity of the developed simulator.

The quantitative differences between experience levels supports the construct validity of the AR wire navigation simulator. Significant univariate performance differences between subject groups were only shown in TAD measurement. It was hypothesized that experienced orthopaedic surgeons would outperform the other subject groups across all dependent variables. The difference in the hypothesized outcome and the experiment outcome, including the additional dependent variables, could be the result of a speed-accuracy tradeoff. Experienced surgeons may carefully ensure accurate placement of the wire, requiring extra time and images. Similarly, inaccurate placement of the guide wire requires less time and fewer images. As previously discussed in Chapter 2, an acceptable TAD is less than 25 millimeters. A TAD above 25 millimeters is shown to have a failure rate of between 27-44 percent [142, 144]. Further review of the results showed that experienced surgeons experienced zero trials with a TAD exceeding 25 millimeters. Six out of 11 (54.5 percent) first-year orthopaedic residents and 13 out of 22 (59.1 percent) undergraduate engineering students had a least one trial with a TAD above 25 millimeters.

Of the 105 trials collected from inexperienced undergraduate students and first-year orthopaedic residents, 33 (31.4 percent) of the individual trials resulted in a TAD above 25 millimeters. Further analysis of these inaccurate trials revealed a trend in the error of guide wire placement. The final, end location of the guide wire tip for inexperienced subjects was frequently located in the posterior hemisphere of the femoral head. Guide wire tip locations in the posterior hemisphere occurred in 78.8 percent of the trials with inaccurate tip-apex distances. Further dissection of the guide wire locations revealed that 30.5 percent of the unacceptable wires were placed in the posterior-inferior zone, a region that has shown to be the most susceptible to mechanical failure after hip fracture repair [142]. A complete dissection of all wire tip locations regarding inaccurate TAD trials is shown in Figure 38.

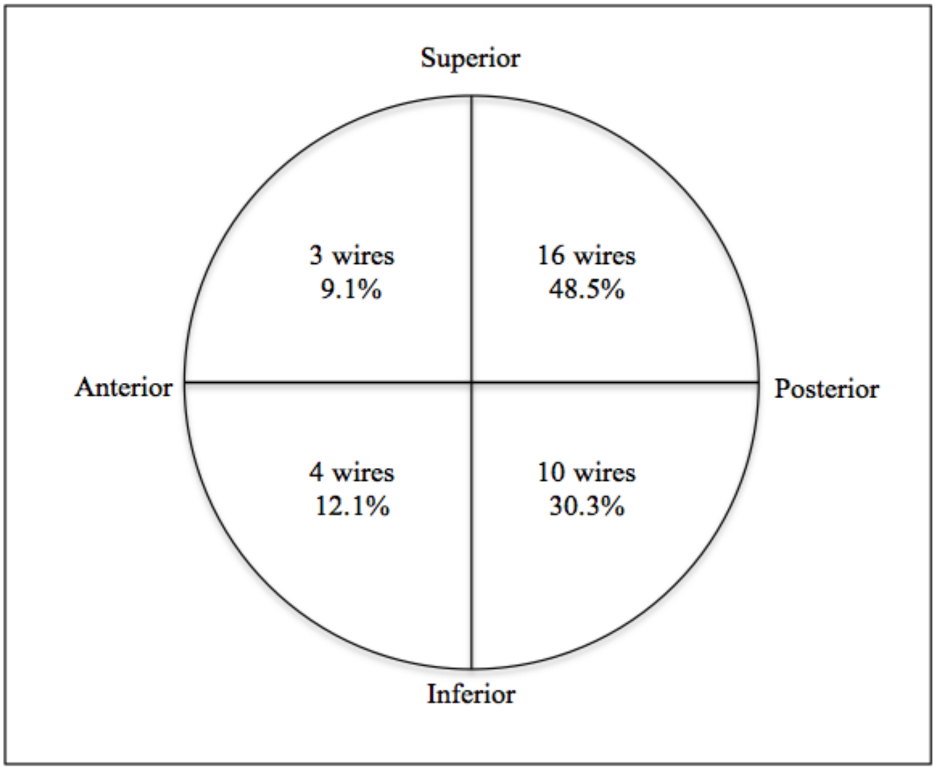

Figure 38. Guide wire tip locations of 33 simulator trials with inaccurate tip-apex distances.

These wire inaccuracy results first suggest that novices do not possess the same level of psychomotor drilling skill as experienced surgeons. In addition, guide wires tend to be inaccurately placed in the posterior hemisphere of the femoral head. This may be the result of inexperienced subjects not accounting for the anteversion of the femoral neck. Not accounting for anteversion of the femoral neck would result in a wire drilled at a horizontal angle, resulting in an ending location within the posterior hemisphere. Another possibility could be the visual difference between the femoral neck angle on the lateral fluoroscope image, compared to the actual angle of the femoral neck. In particular, the lateral fluoroscope image shows the femoral neck at a horizontal angle, which could be deceiving to an untrained subject.

In addition to inaccurate guide wire placement, undergraduate engineering students made significantly fewer attempts to place the wire compared to first-year orthopaedic residents. Although engineering students required fewer drilling attempts than first-year orthopaedic residents, this group was also the least accurate in the final TAD.

There were no differences between accuracy of the entry location of the guide wire. This suggests that the differences in performance occur after finding the correct starting location for drilling; further suggesting that the difficulty of aiming and drilling the guide wire to the apex of the femur is the difficult skill of the task.

However, the study does have several limitations. The sample size of participants with an orthopaedic surgical background is relatively small, which restricts the generality of the validity study, particularly due to the difficulty of securing orthopaedic surgeons' time and the limited number of orthopaedic surgical trainees at a single medical institution. However, the results reported are unique and serve as a reference point for future simulator development in wire navigation. Future trials may need to revise the experimental design for trials not achieving a satisfactory TAD less than 25 millimeters for the possibilities of viewing univariate effects across more dependent variables. In addition, this trial places the same weight across all dependent variables, where one may hypothetically place more importance on accuracy and number of drilling attempts, compared to the number of images used and the duration of the trial.

CHAPTER 5 – AIM THREE: SKILL TRANSFER FROM SIMULATOR

5.1 Introduction

The third aim is to provide evidence to support the concept that simulator training promotes skill transfer to a more realistic scenario by using real fluoroscopy. This section details trials of assessing first-year orthopaedic residents after performing training on the developed augmented reality wire navigation simulator. The assessment uses a real C-arm fluoroscope operated by a radiology technician. In addition, wire navigation assessment is completed on a soft-tissue hip model (Sawbones® part number 1516-3) equipped with replaceable radiopaque femurs (Sawbones® part number 1130-177). The assessment setup is shown in Figure 39. These experimental trials are initial steps in proving the training on developed simulator transfers skill to a real operating room.

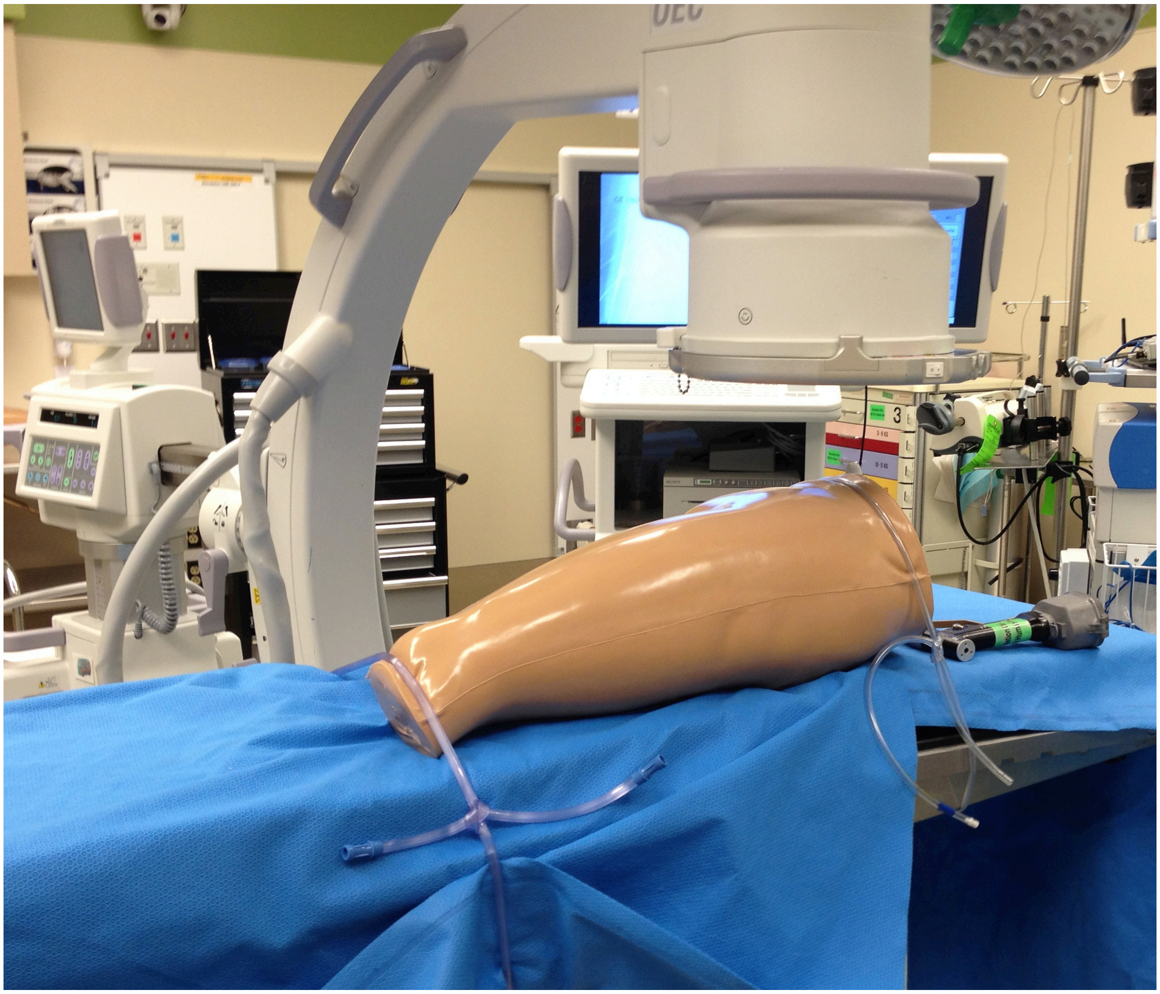

Figure 39. Wire navigation assessment device using real fluoroscopy and a soft-tissue artificial hip model.

5.2 Hypothesis

Orthopaedic residents trained on the augmented reality wire navigation simulator will show superior performance compared to residents not trained on the simulator. The dependent variables for assessing performance include wire accuracy (i.e., TAD), duration of wire navigation, number of fluoroscopic (x-ray) images used, and the number of drilling attempts.

5.3 Methods

5.3.1 Participants

Eight subjects participated in the study from the University of Minnesota. Both the University of Iowa and University of Minnesota Institutional Review Boards approved the study. All subjects were first-year orthopaedic residents who participated as part of a skill development day where students received lectures in topics such as fluoroscopy, wire fixation, and compartment syndrome. The eight subjects were divided into two equal groups: the trained group and the untrained group. All subjects received 40 minutes of lecture instruction on wire navigation and fluoroscopy before starting the experiment. In addition, all subjects were given basic instructions on the experiment explaining fluoroscopy guidelines and assessment metrics before beginning wire navigation on the artificial hip. However, the trained group received instruction and coaching on the augmented reality wire navigation simulator before being assessed using real fluoroscopy whereas the untrained group did not. All participants completed two consecutive wire navigation trials on the artificial hip model using real fluoroscopy.

5.3.2 Augmented Reality Wire Navigation Simulator Training

The trained group performed a training program on the developed augmented reality wire navigation simulator before assessment using real fluoroscopy. Prior to starting the session each subject was given an explanation of the technique for wire navigation, which was given approval by faculty orthopaedic surgeons. The training session was observed and coached by a faculty orthopaedic surgeon. The wire navigation procedural technique is shown in Table 4.

Table 4. Directions for wire navigation on the augmented reality simulator.

1. Palpate the femur to find the lateral surface.
2. Place the wire on the lateral surface at the approximate starting point. The starting point will be a few centimeters distal to the vastus ridge and proximal to the lesser trochanter. The starting point should be halfway in between the anterior and posterior cortex.
3. Horizontally adjust the angle of the wire using the AP view. Ensure the projected trajectory of the wire intersects with the apex of the femoral head.
4. Drill the guide wire a few millimeters into the lateral cortex.
5. While still penetrating the lateral cortex, check the lateral view and angle the guide wire so the projected trajectory of the wire intersects the apex of the femoral head.
6. Drill the wire into the femur several centimeters and check the progress in both the AP and lateral views. If the progress is satisfactory, proceed with the drilling, if not, back the wire out the femur and start over.
7. Drill and check the progress in the AP or lateral view. Make fine drilling adjustments and ensure the tip-apex distance is less than 25 millimeters.
8. Before finishing, check both the AP and lateral view to ensure the wire has not penetrated the joint and the wire placement is satisfactory.

The subjects individually performed four trials on the simulator for training. The first trial on the augmented reality wire navigation simulator overlaid the ideal trajectory for the wire on the virtual fluoroscopy. This overlay showed both the ideal starting location and finishing placement. This was intended as a tutorial for correctly inserting the wire. An example of the virtual fluoroscopy overlay showing the ideal trajectory is shown in Figure 40. The following three training trials did not include the overlay.

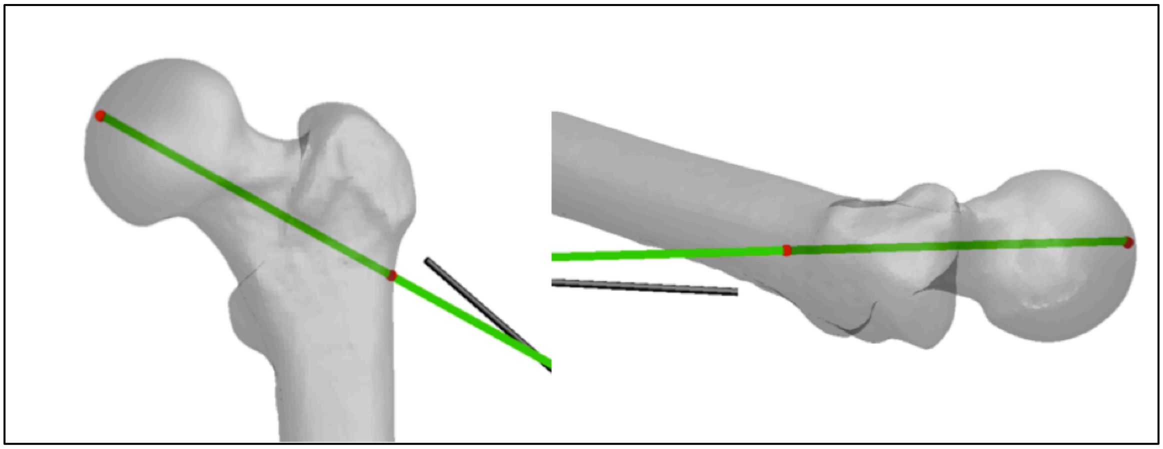

Figure 40. The virtual fluoroscopy of the augmented reality wire navigation simulator showing overlaid trajectory for teaching.

In addition to onscreen instruction, subjects were coached for further developing individual wire navigation skills. The faculty surgeon observed weaknesses of each subject's technique and offered constructive direction. The surgeon also answered questions from subjects regarding the specific wire navigation technique. Instructions included verbal coaching cues such as minimizing drilling forces by using high spindle speeds and resting the secondary hand on the table for drilling stability.

5.3.3 Performance Assessment using Real Fluoroscopy

Both the trained and untrained groups underwent performance assessment using real fluoroscopy on a soft tissue hip model. A subject performing the wire navigation assessment is shown in Figure 41. Each subject repeated the assessment two consecutive times. Fluoroscope images were saved for calculating and assessing the TAD. In addition to TAD, subjects were assessed on duration to complete the wire navigation, number of drilling attempts, and number of fluoroscopy images used through each trial.

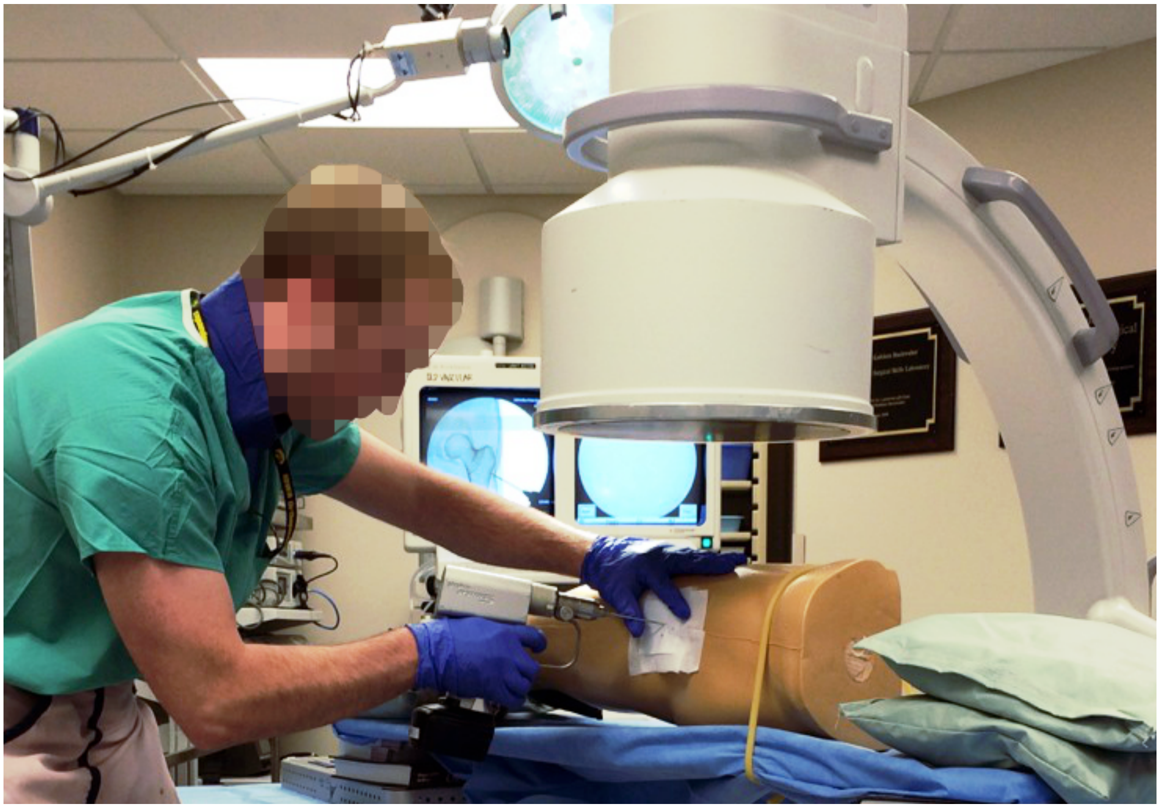

Figure 41. Subject performing wire navigation using real fluoroscopy on soft-tissue hip model.

A scoring system was set for placing procedural tradeoffs between different dependent variables. The scoring system was a consensus between engineers and a faculty surgeon. The goal of the scoring was achieving the lowest score possible. Every second during the trial was assessed as one point. Each image used throughout the trial was also assessed as one point. Every millimeter of error in the TAD was assessed as ten points. Reversing the wire out of the femur and drilling again was assessed as a fifty-point penalty for every successive breach of the lateral cortex. In addition, a TAD above the acceptable 25 millimeters was considered a failed trial. A summary of the scoring system is shown in Table 5. Performance differences between training groups were assessed using a repeated measures multivariate analysis of variance (MANOVA).

Table 5. Scoring system for assessment of wire navigation using real fluoroscopy.

Category	**Points**
Tip-apex distance	10 points per millimeter of error
Duration	1 point per second
Fluoroscopic images	1 point per image
Drilling attempts	50 points per breach of the lateral cortex

5.4 Results

Completing a repeated measures MANOVA for all subjects (including the failed attempts) showed non-significant main effects for training, revealing Wilks' Lambda = 0.466, F(4.0, 3.0) = 0.860, p = 0.573, partial eta squared = 0.534. Observed power was 0.111. The univariate results are shown in Table 6.

Both the multivariate and univariate analysis reveal low observed power for the experimental results. Three of the eight subjects obtained a TAD above the acceptable 25 millimeters. Two of these subjects were in the trained group.

Table 6. Non-significant univariate effects between training groups.

Measure	F	Significance (p)	Partial Eta Squared	Observed Power
TAD	1.317	0.295	0.180	0.163
Fluoroscopy Images	1.918	0.215	0.242	0.216
Duration	2.184	0.190	0.267	0.239
Drilling Attempts	1.000	0.356	0.143	0.136

5.5 Theories for Non-significant Results

Moving surgical training to simulation outside the operating room is a crucial step to improving safety in the current training model. This research makes an initial step towards showing a developed augmented reality wire navigation simulator that improves skill in a more realistic scenario using real fluoroscopy. The long-term objective of this research is a proven simulator for developing wire navigation skill for hip fracture surgery. This study compares wire navigation assessments of orthopaedic surgical residents using real fluoroscopy between two groups: one group trained on the augmented reality wire navigation simulator, and the other group receiving no hands-on training before assessment.

The results exhibit an unexpected result that does not show a significant performance difference between first-year residents trained on the simulator and untrained residents. This raises several important questions: Aside from sample size, why was there no observable difference between the training groups? Also, why did three of the eight assessments for the trained group result in failed outcomes due to wire inaccuracy? After further analyzing the data, I propose four possible theories for the observed results:

- A more extensive training program is needed to quantify and observe improvement from repetitions on the augmented reality wire navigation simulator.
- Novice subjects do not possess the skill to objectively and accurately estimate tip-apex distances during the wire navigation task.
- Although there is a high variability of initial skill between the observed subjects, each subject may be making cognitive decisions regarding the speed-accuracy trade off of the wire navigation task.
- There exists some discrepancy between the wire navigation task on the augmented reality simulator and the assessment using real fluoroscopy.

The following examines each one of these theories individually. This preliminary investigation individually analyzes results specific to each of these proposed theories.

5.5.1 Extended Wire Navigation Training

Wire navigation is a complex psychomotor skill that may longer than to develop than the provided training session that included four training trials. Bjorgul et al. observed a learning period of twenty to thirty orthopaedic hip surgeries using cannulated screws before statistically significant improvement was detected [54], shown in Figure 42. This suggests that extending the training period over several days would produce quantifiable improvements in the wire navigation task. The question remains, how much practice produces significant improvement?

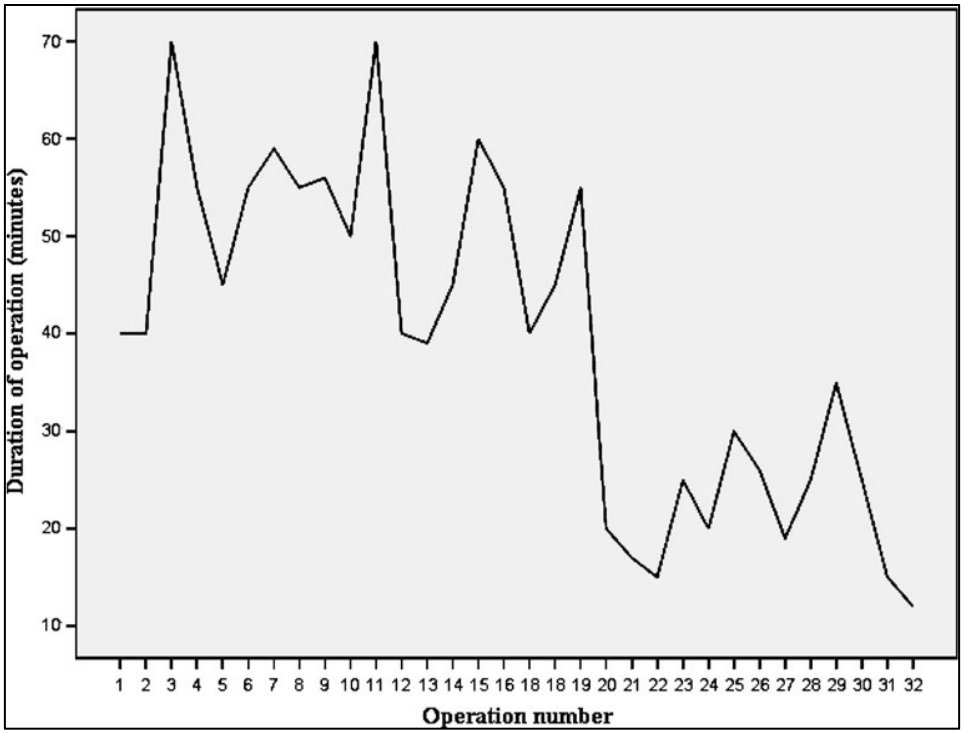

Figure 42. Duration of surgery improvement using cannulated screws for one resident from [54].

Significant improvement is difficult to predict due to the various training conditions of the task. The acquisition of skill is dependent on many factors, including initial skill before training, the individual's ability to learn, and the quality of coaching/feedback throughout training. All of these factors contribute to orthopaedic drilling skill acquisition and the slope of the learning curve. The constructed augmented reality wire navigation simulator enables training on a non-displaced fracture that provides feedback on the wire accuracy, time to complete the drilling, and number of imaged used throughout the wire placement. This training allows the construction of a baseline learning curve with minimal coaching and feedback. It is expected with expert coaching and feedback the skill acquisition would be faster for orthopaedic wire navigation. This baseline curve could be used for comparative analysis in the future to assess the effectiveness of surgical coaching.

To construct a baseline learning curve for the wire navigation task, five engineering students were recruited to perform wire navigation training on the augmented reality simulator over four days. Each day consisted of eight trials and a maximum of one hour of exposure to the augmented reality wire navigation simulator. Feedback was given between each successive trial consisting of the TAD, time to complete the trial, and number of images used throughout the trial. The 32 trials established a learning curve for each subject on the wire navigation simulator. Linear regression was completed to quantify the effects of practice in wire navigation performance.

A linear regression analysis was conducted to determine relationship between the wire accuracy (i.e., TAD) and the number of practice trials. The average TAD learning curve for all subjects is shown in Figure 43. There was a significant relationship between the TAD and trial number (Pearson's $r = -0.344$, $p < 0.001$). The regression plot is shown in Figure 44. The regression equation is,

$$\text{TAD} = 15.8 - 0.21 * \text{Trial Number} \qquad (7)$$

The number of practice trials significantly predicts the wire accuracy, F(1,158) = 21.2, $p < 0.001$. The result demonstrates the number of practice trials predicted 11.8% of the variance in TAD. The standardized slope (-0.3) and the unstandardized slope (-0.2) show a significant difference from zero ($t = -4.6$, $p < 0.001$), indicating a training effect. The TAD is predicted to improve 0.1 to 0.3 millimeters with each successive practice trial (95% CI, $p < 0.001$). This suggests between three and ten trials are needed to improve one millimeter in the TAD, revealing wire navigation accuracy is a slow developing skill.

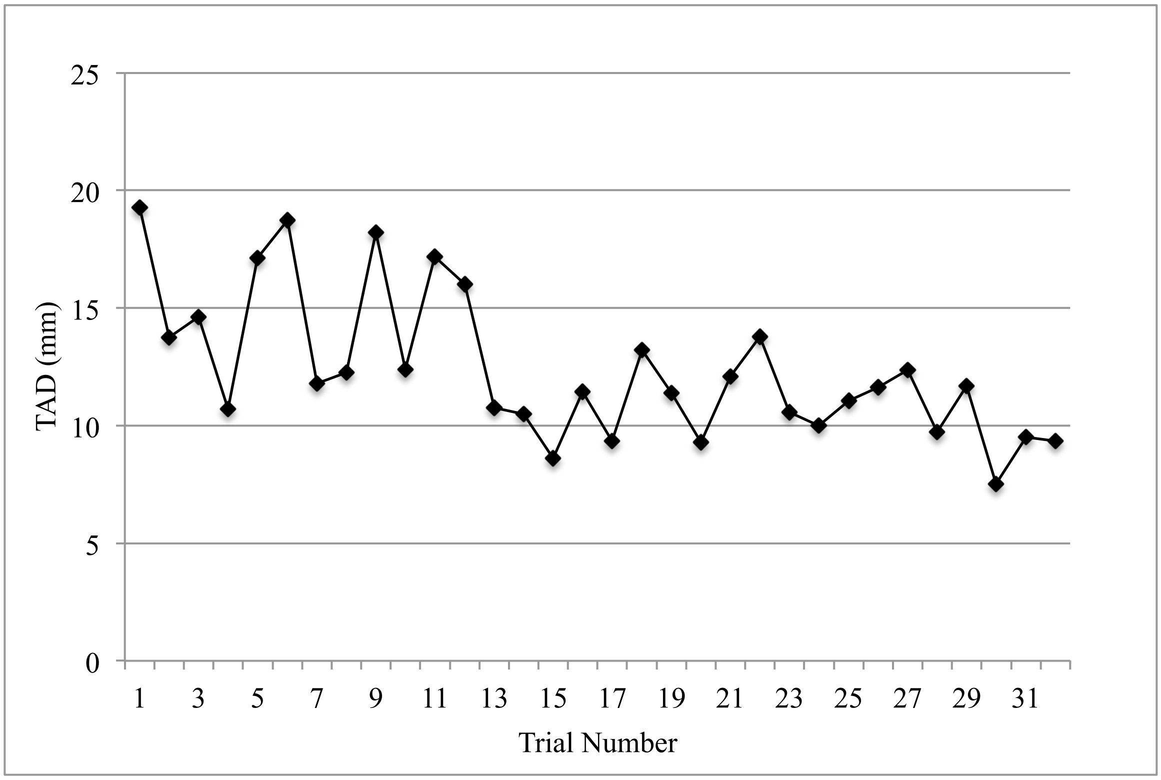

Figure 43. The average TAD per trial for five engineering students on the augmented reality wire navigation simulator.

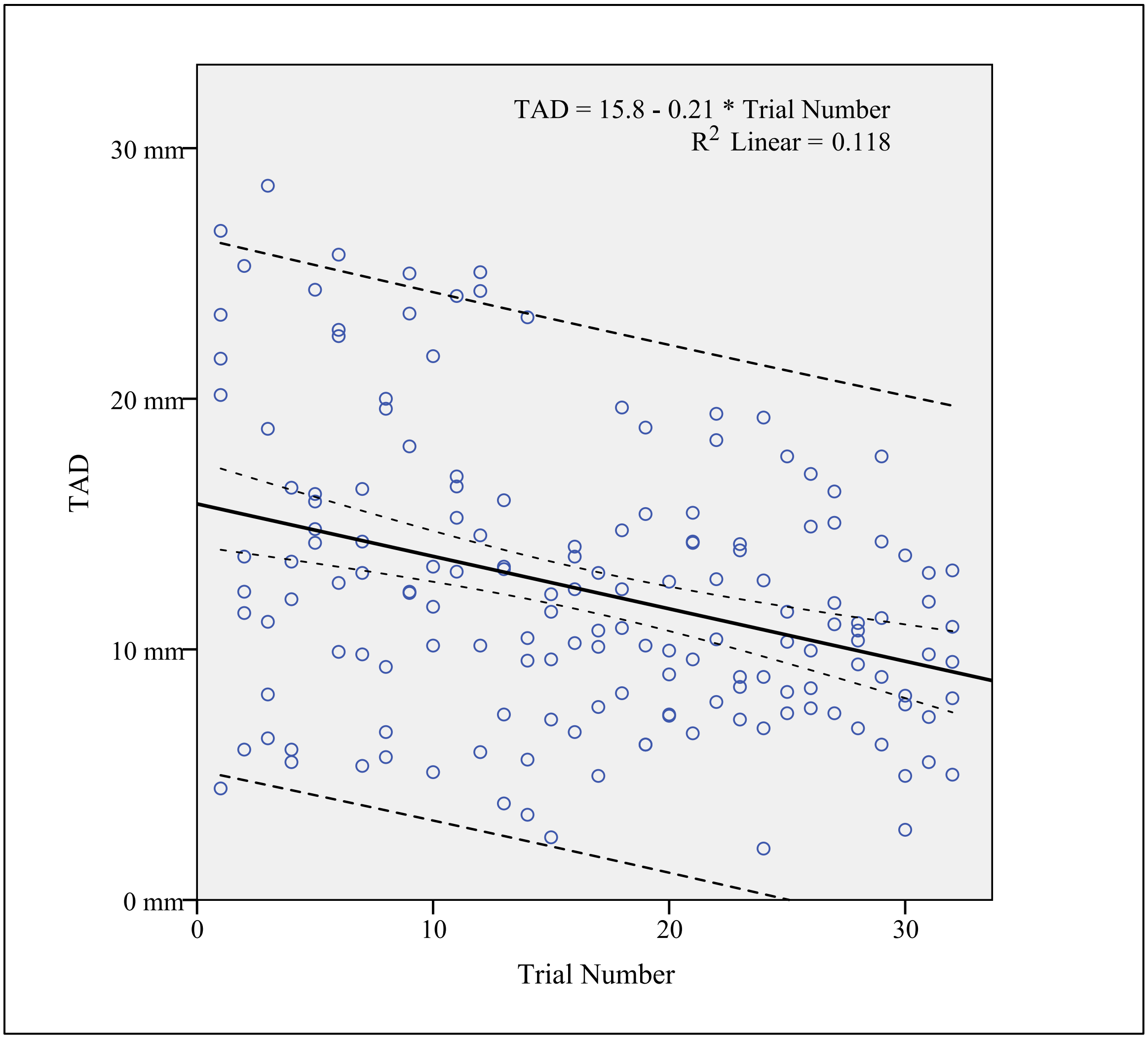

Figure 44. The TAD vs. practice trial number. The dashed lines show the 95% individual prediction interval and mean prediction interval.

In addition to wire accuracy, the number of practice trials significantly predicted improvement in time to complete wire navigation, F(1,158) = 9.11, p = 0.003. The regression equation is,

$$\text{Duration} = 107 - 1.2 * \text{Trial Number} \tag{8}$$

The average duration per trial is shown in Figure 45. The standardized slope (-0.2) and the unstandardized slope (-1.2) show a significant difference from zero ($t = -3.0$, $p = 0.003$), indicating an improvement over the repetitions. The time to complete wire navigation improves between 0.4 and 2.0 seconds each subsequent practice trial (95% CI, $p < 0.001$). The regression plot is shown in Figure 46. This shows that subjects complete the task 30-seconds faster after approximately 25 additional practice trials on the simulator. Both regression analyses support the theory that extended practice is required to develop skill. It may be possible to expedite learning through the development and integration of effective coaching and feedback techniques on the simulator. This supports the hypothesis, showing future experiments need to extend training over several days and/or integrate more effective coaching methods to detect quantifiable benefits of the wire navigation simulator.

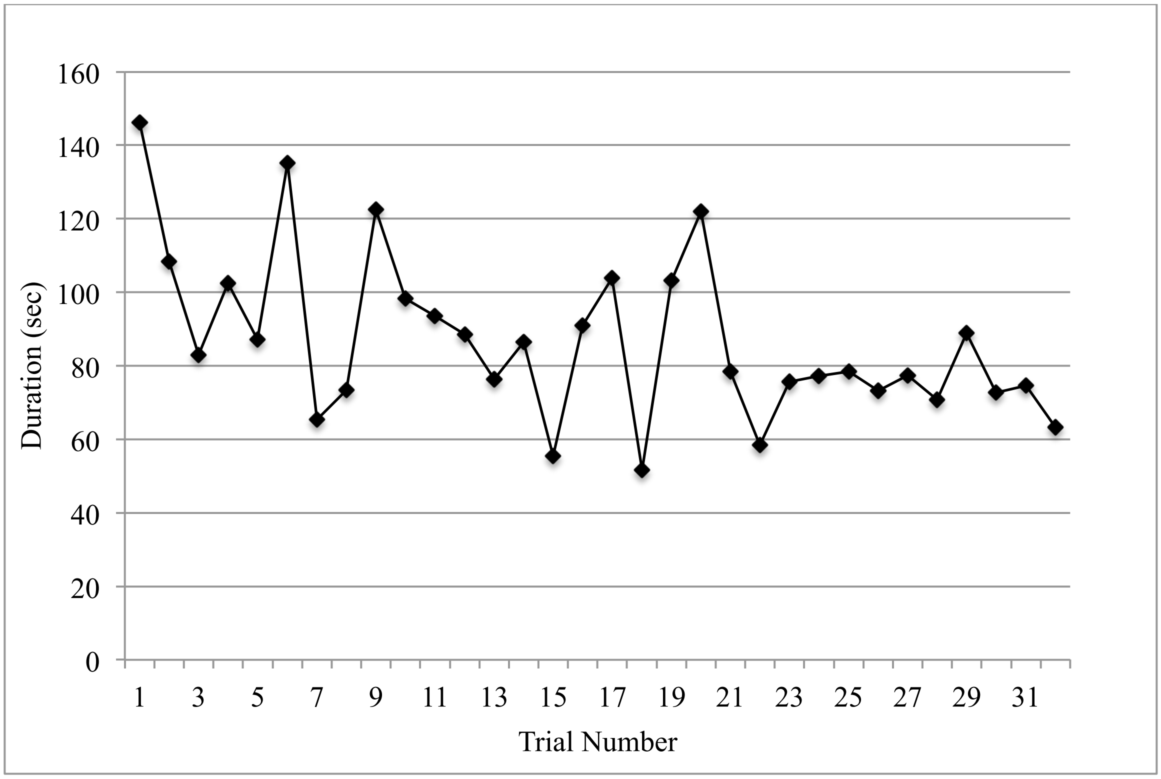

Figure 45. The average duration per trial for five engineering students on the augmented reality wire navigation simulator.

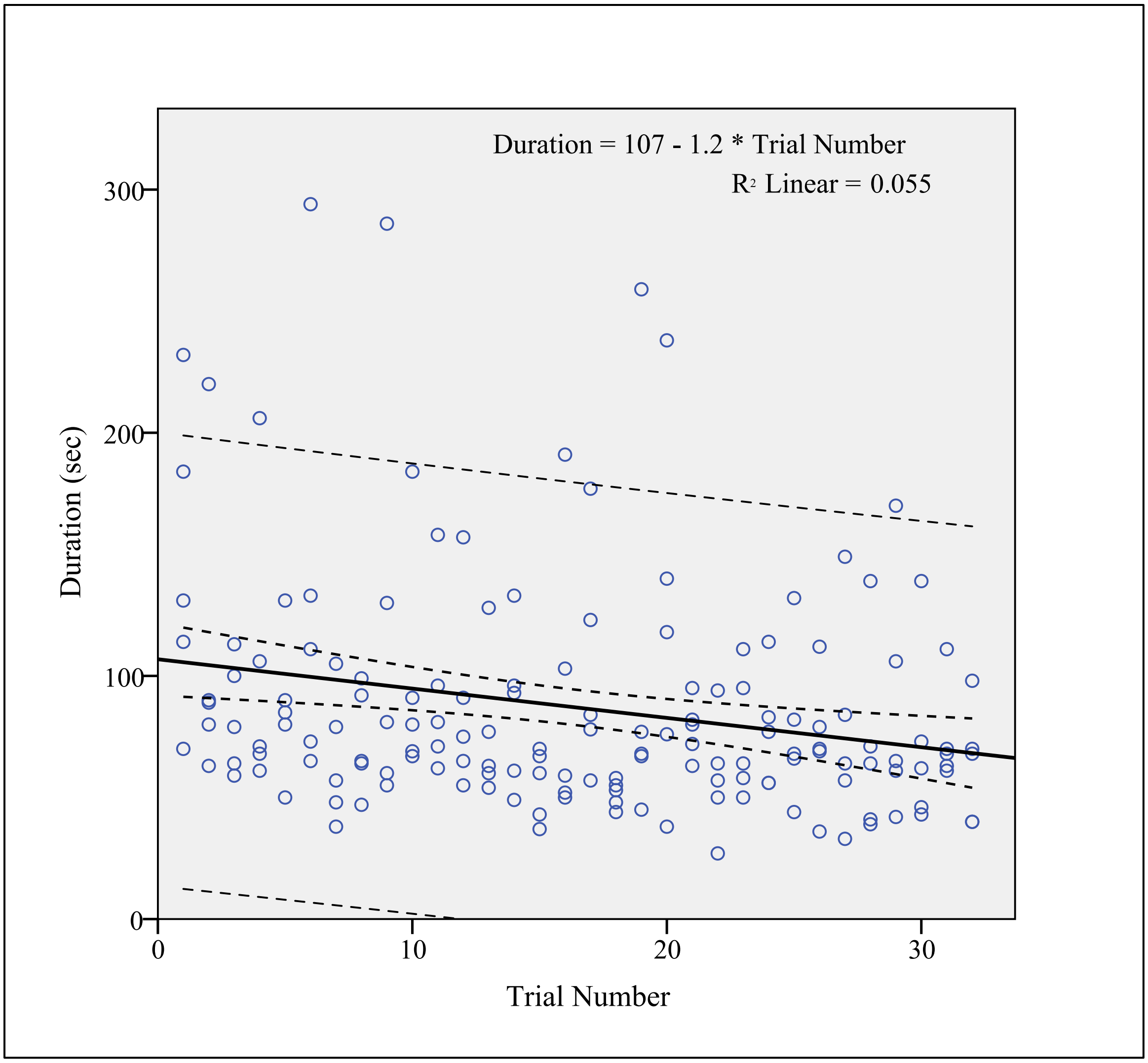

Figure 46. Duration of the wire navigation trial vs. the practice trial number. The dashed lines show the 95% individual prediction interval and mean prediction interval.

5.5.2 Observational Accuracy of the Tip-Apex Distance

Subjects were allowed unlimited drilling attempts to achieve an accurate TAD in the assessment using real fluoroscopy. Even with unlimited drilling attempts, three out of the eight subjects achieved a TAD above 25 millimeters on the assessment. These results suggest novice subjects cannot objectively estimate the TAD, which is a prerequisite skill to accurately drilling a wire. This miscalculation of accuracy could be detrimental and lead to more mechanical failures of hip implants.

To test this theory an online survey was designed. The online survey consisted of 16 sets of radiographic images with each set showing an AP and lateral view. The participants were instructed to estimate the TAD for each set of radiographic images. Each participant was given 30 seconds per set of images. This time limit ensures it is TAD estimation, not an exact calculation. The subjects were given no feedback. The survey was administered to first-year orthopaedic residents at the University of Iowa who had previously completed wire navigation training, which included an explanation of how to estimate TAD from fluoroscopic images.

Five subjects finished the survey. The average error for the 16 sets of radiographic images was 12.4 millimeters with a standard deviation of 12.0 millimeters ($\mu = 12.4$ mm, $\sigma = 12.0$ mm). This highlights the lack of a critical skill for wire navigation. Without accurate estimation, the physical drilling skill is irrelevant. This demonstrates novices cannot accurately estimate distances in radiographic images, supporting this hypothesis.

5.5.3 Speed-Accuracy Tradeoff

The concept of the accuracy decreasing with increasing speed was first proposed by Fitts [153], advocating the performance capacity is associated with the visual and proprioception feedback loops. The wire navigation assessment showed a wide range of speed and accuracy. The time to complete the task ranged from one minute and fifteen seconds to four minutes and twenty seconds. The wire accuracy ranged from tip-apex distances from 8.3 millimeters to 41.5 millimeters. Although there was a wide range of initial skill levels, the individuals exhibited speed-accuracy tradeoffs between their two trials. Seven out of eight subjects remained constant or improved their accuracy by taking more time to complete the task. This is graphically shown in Figure 47. This supports the notion that although skills vary in speed and accuracy, each subject makes decisions, whether conscious or unconscious, about the tradeoff between the speed and accuracy. This suggests, independent of initial skill levels, novice residents that take more time to setup and drill the wire finish with a more accurate result. The scoring system established for the wire navigation task was intended to explicitly define tradeoffs between accuracy and speed. A possible reason this scoring system was ineffective was novice subjects were too focused on the task and unaware of the exact elapsed time. Another possible reason for the ineffectiveness of the scoring system is the subjects' inability to estimate TAD resulting in an inability to estimate their true score.

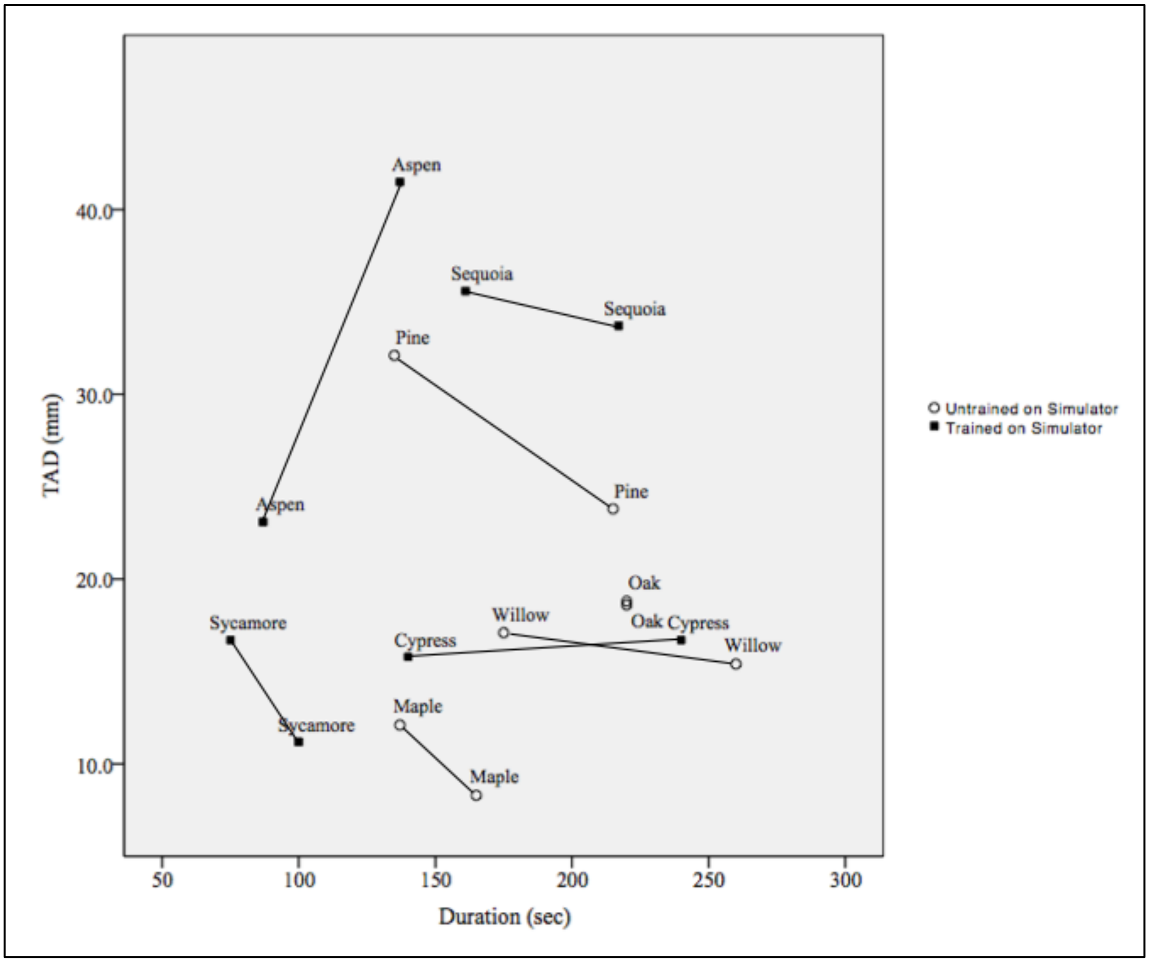

Figure 47. Individual tradeoffs between wire accuracy (TAD) and the duration to place the wire.

5.5.4 Discrepancies between the Augmented Reality Simulator and using real Fluoroscopy

The lack of training effect may be caused by incorrectly simulating important details in the wire navigation task between the assessment and training. One probable discrepancy is the orientation of the femur on the simulator. The augmented reality simulator places the femur at a ten to fifteen-degree inclined anteversion of the femoral neck, similar to angles in real scenarios. In the simulator training, subjects are instructed to "drop the hand" to account for the anteversion of the femoral neck. Several subjects stated that after completing the assessment using real fluoroscopy, the inclination of the femoral neck in the assessment was less than in the training simulator. This is likely because novices followed closely to procedural steps learned during training and did not possess the skill to adapt to unfamiliar scenarios. The training likely caused subjects to overcompensate for the anteversion and drop the drilling hand too much. This overcompensation resulted in a majority of wires on the anterior side of the femoral head. Seven out of the eight trials for the subjects trained on the augmented reality simulator ended in the anterior region of the femoral head. This is the opposite effect seen in untrained novices assessed on the augmented reality simulator in Chapter 4. These novices, not accounting for the inclination angle due to the anteversion of the femoral neck, primarily drilled the wire in the posterior region of the femoral head. This shows coaching to "drop the hand" could train an erroneous behavior in the novice participants when inclination angles are slight or nonexistent.

5.6 Discussion

Although this study does not provide evidence simulator training transfers to a more realistic scenario using real fluoroscopy, several more interesting discoveries are made. These additional findings further advance the understanding of the wire navigation task. First, the training curve experiment demonstrates that wire navigation is a slowly developing skill that takes several days of practice before improvement is observable. Roughly five practice trials are required to show improvement of one millimeter in the TAD and the wire navigation speed improves roughly one second with each repetitive practice trial. The survey showed novices often could not accurately estimate the TAD from radiographic imagery. Presumably this skill should be trained before proceeding to actual drilling trials. Although novices widely vary in initial skill a relationship was revealed between the time to complete the wire navigation task and the accuracy of the wire. This finding shows taking more time results in a more accurately placed wire. All of these findings reveal new discoveries in the orthopaedic field, which can be used for more effective training and simulator development.

5.7 Future Work and Limitations

The main limitation of this study is the sample size and underpowered results. Due to the small resident numbers at any given institution, a multi-institution experiment would be necessary to observe training effects from the augmented reality wire navigation simulator. The result of the power analysis yielded an estimated sample size of 36 subjects using effect size estimations from simulator trials. There are generally only one or two orthopaedic training programs in a single state. However, it is possible with cooperation between multiple institutions. In addition, the learning curves for the wire navigation were performed with engineering students with minimal coaching. This scenario limits the generalizability to surgical residents that receive coaching throughout the training trials.

CHAPTER 6 – DISCUSSION

6.1 Discussion of Aim One: Simulator Design & Development

The first aim is shows augmented reality can be effectively used to develop a wire navigation simulator. This aim proposes a substitute method for developing simulators using stimulus-response cues of the task when complete replication of the actual task is not possible. This research is not the first to explore orthopaedic drilling in a simulation environment for hip fractures, but it is original in its application of technology.

Past simulators have existed in two categories: 1) software-based simulation with no drilling forces [95, 96, 98, 99, 154, 155] and 2) devices using haptic force feedback devices for drilling recreation [1, 2, 100, 101]. This research is the first successful attempt to implement augmented reality (AR) into an orthopaedic wire navigation simulator. This AR wire navigation simulator uses a real surgical drill and is augmented by virtual, radiation-free fluoroscopy. The virtual fluoroscopy is made possible by electromagnetic sensors tracking the position of the femur and drill simultaneously. These unobtrusive electromagnetic sensors are another unique advantage over orthopaedic drilling simulators using Phantom® force feedback devices for motion tracking. As a result of this unique approach, the developed AR wire navigation simulator contains important sensory cues that are overlooked and unavailable in other simulators. In addition, calibration algorithms are developed and documented in the AR navigation simulator, which can be used by future simulator developers to pursue further advancement in augmented reality technology.

Although the AR wire navigation simulator contains training features not available in real surgery, limitations exist in the simulator. The first limitation is the electromagnetic nature of the motion tracking. This system is both relatively expensive (~$10,000) and is susceptible to metallic interference. It is imperative to eliminate all unnecessary metallic components from the simulator environment for the simulator to function properly. The surgical drill contains substantial amounts of metal, which has minimally affected the accuracy of motion tracking. In addition to metallic interference, the flexibility of the wire has continually been a source of inaccuracy in the simulator.

6.2 Discussion of Aim Two: Construct Validity of the AR Wire Navigation Simulator

The second aim of this study examines the construct validity of the augmented reality wire navigation simulator. This aim demonstrates the simulator is successfully designed to measure and differentiate between the traits of novice and experienced surgeons. A valid training and assessment device is a way to demonstrate the simulator's ability to measure and report traits between levels of surgical skill. This aim completes the most extensive construct validation study of any orthopaedic drilling simulator to date. In addition, this is the first study to differentiate drilling accuracy between novice and experienced surgeons. This hints the augmented reality could offer superior benefits to the force feedback devices.

Forty subjects were separated into three groups according to their level of surgical experience: undergraduate engineering students, first-year orthopaedic residents, and experienced surgeons. The experienced group consisted of forth-year orthopaedic residents and faculty orthopaedic surgeons. The experienced group was predicted to outperform the novices across all dependent variables including duration to complete a wire navigation trial, number of images used during the trial, drilling attempts, and wire positional accuracy.

The observed results show the simulator is able to differentiate experience levels over the measured dependent variables (F(10,70) = 2.375, p = 0.017). Although experienced surgeons took less time to complete the wire navigation, the only significant difference in performance was in the TAD (F(2,39) = 3.721, p = 0.033). The time required to complete the wire navigation was approaching significance with p = 0.066, observed power = 0.536. The time and accuracy of the wire were the two major components for detecting differences in skill in the wire navigation task. A comparison of accuracy and time on the simulator between skill levels is shown in Figure 48.

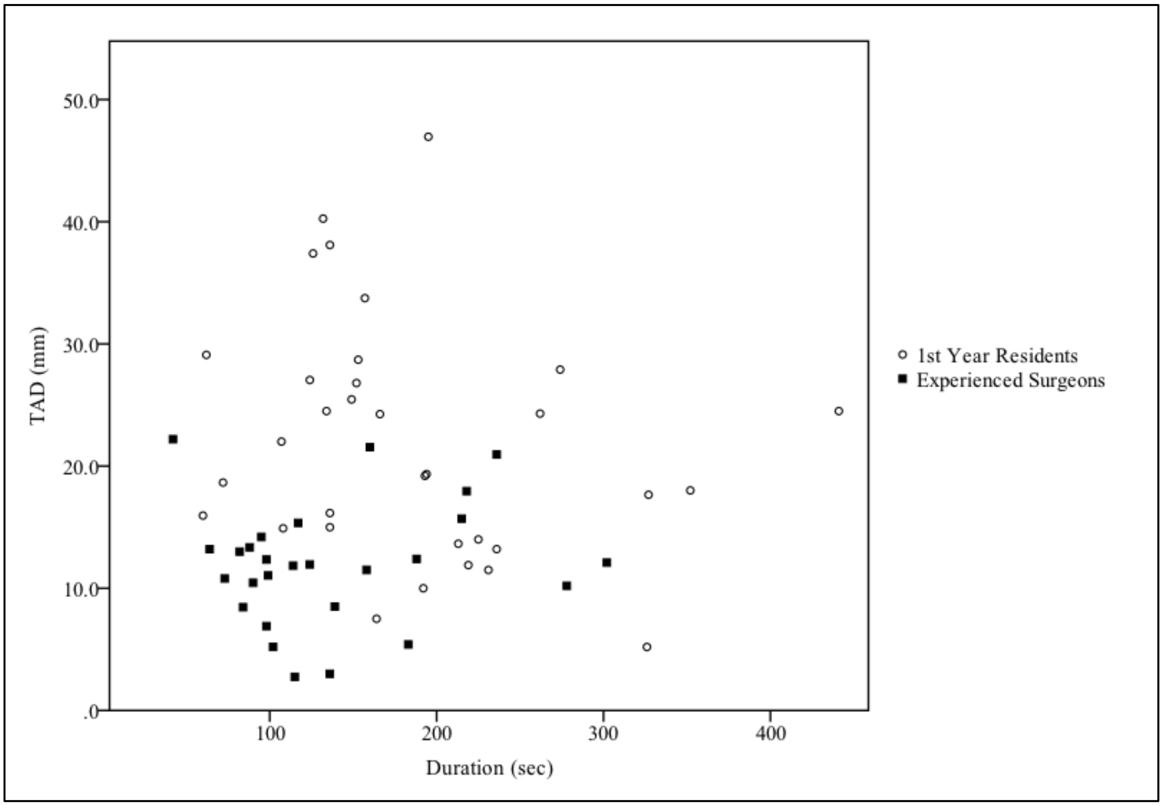

Figure 48. Comparison of accuracy and time between experienced surgeons and first-year residents on the wire navigation simulator.

Undergraduate engineering students took fewer drilling attempts than participants in the more experienced groups. As a result of making fewer drilling attempts, undergraduate students were achieved the least accurate wire placement of all groups. Froelich et al. observed the same effect in orthopaedic drilling, stating “We believe this finding may be a reflection of either the junior resident’s willingness to accept or inability to recognize a less than optimal starting point in comparison with a more senior resident" [1].

In addition to the measured dependent variables, this study identified another characteristic behavior in wire navigation by novices. Novices tend to error on the posterior side of the femoral head with the wire. This is the result of novices not realizing or adjusting for the anteversion of the femoral neck. The recorded positions show novices accommodate for the angle of the femoral neck in the frontal plane, but often neglect the fifteen-degree anteversion angle of the neck in the horizontal plane. The identification of this previously unknown behavior potentially enables more effective coaching techniques of the wire navigation task.

6.3 Discussion of Aim Three: Skill Transfer from Simulator

The initial goal of aim three was to demonstrate that practice on the simulator improves performance in a scenario using real fluoroscopy. Two subjects trained on the augmented reality wire navigation simulator with expert coaching failed the assessment by attaining tip-apex distances exceeding 25 millimeters. Although training on the simulator did not effectively improve performance on a more realistic scenario, several key discoveries were made through additional investigation.

The additional investigation discovered skill acquisition occurs much more slowly in wire navigation than previously assumed. Noticeable improvement may take several days of repetitive practice on the wire navigation task. This demonstrates many of the orthopaedic trainees may not receive enough exposure to the wire navigation task throughout residency to develop proficiency. This demonstrates the need for supplemental wire navigation simulator training outside the operating room.

Another interesting characteristic these suboptimal results showed is the novices' inability to estimate an inaccurately drilled wire in the proximal femur using fluoroscopy.

In one-fourth of the trials, the subjects achieved a TAD above 25 millimeters. The subjects achieving these results should have reversed the wire and reattempted the wire navigation, but failed to accurately estimate the accuracy of the wire. To quantify the accuracy of novices' estimation ability, novices were surveyed using 16 fluoroscopy image sets, which contained both an AP view and a lateral view for estimating TAD. The results demonstrate that novices lack the ability to accurately estimate distances on radiographic images. This shows a lacking skill which can be developed outside the operating room which would further improve performance in the wire navigation task.

The real fluoroscopy assessment revealed a speed-accuracy tradeoff in each subject. Although individual skills on the wire navigation task greatly vary, the trend shows subjects who take longer to complete the wire navigation task achieve a better wire accuracy. The accuracy of seven out of eight novice subjects remained consistent or improved when the subjects took more time to complete the wire navigation of the proximal femur. This advocates the notion that novices should primarily practice drilling a wire accurately before attempting to maximize their speed or minimize their number of fluoroscopic images. Once novices consistently achieve wire accuracy, more emphasis can be placed on efficiency.

The training and assessment of the experiment made the researchers aware of discrepancies between the augmented reality simulator training and assessment. One of the most notable discrepancies is the angle of anteversion of the femoral neck. An expert surgeon operates on a hip with varying angles in the frontal plane as well as varying anteversion angles. An expert surgeon is also able to adapt to varying angles and makes adjustments in accordance to fluoroscopy. However, novices trained on a simulator using a single proximal femur model may develop technique bias to that single case. This is observed in the differences between the augmented reality training and assessment using real fluoroscopy. Trainees that were trained on a fifteen-degree anteversion of the femoral neck were unprepared for an anteversion neck less than fifteen degrees. These novices used a more procedural-based technique to complete the task. The procedural steps varied when “dropping the hand” was not needed in the assessment. As a result, seven out of eight trials by the subjects trained on the augmented reality simulator ended with the wire in the anterior region of the femoral head.

CHAPTER 7 - CONCLUSION

The goal of this research is to better understand skill development for orthopaedic drilling using a validated simulator. Effective development of an augmented reality wire navigation simulator enabled researchers to discover and examine important characteristics of orthopaedic wire navigation. The hardware and calibration techniques resulted in a system with sufficient accuracy to convincingly and reliably reproduce the critical stimulus and response characteristics of the wire navigation task. This work also defines techniques that differentiated the skill of experienced surgeons and novices completing the wire navigation task, an unexplored area in the orthopaedic field. This was the first orthopaedic study to show statistical evidence of experienced surgeons outperforming novices in orthopaedic drilling accuracy on a simulator.

The study effectively developed an effective augmented reality wire navigation simulator for differentiating skill level. However, the study failed to prove that training on the simulator directly transfers to more realistic drilling tasks. Further investigation of this unexpected result revealed several new findings related to skill development in orthopaedic wire navigation. For example, skill acquisition occurs much more slowly in wire navigation than previously assumed. As such, a novice is unlikely to become proficient in the wire navigation task over a single day of training. This data shows that navigation skills are gradually improved overextended repetitions of the wire navigation task. This further advocates the need for wire navigation simulators in surgical training curricula.

The developed simulator also identified several characteristic flaws in novice technique. First, novices do not account for the anteversion of the femoral neck resulting in wires placed in the posterior side of the femoral head. This previously unknown behavior potentially reveals that novices have an underdeveloped mental model of the anatomy of the proximal femur. Similarly, novices who were coached to “drop the drilling hand" still navigate the wire inaccurately, suggesting novices have difficulty interpreting the inclination from radiographic images. Another characteristic challenge for novice orthopaedic residents is accurately estimating distances, including the TAD, from radiographic images. Novices incorrectly estimated the wire accuracy by an average of 12.4 millimeters. Tip-apex distance estimation is a foundational skill for wire navigation in the femoral head. Without this skill, the novice cannot assess the final wire placement. This skill could possibly be taught and practiced outside the operating room, which would likely result in better wire accuracy estimation.

Beyond these contributions, these results have several potential practical benefits to the field of orthopaedics. First, a practical and effective simulator was developed to train novice orthopaedic residents in the wire navigation task. This device has already been used at two orthopaedic surgical institutions for surgical skill training days. This augmented reality wire navigation simulator gives trainees an effective, risk-free practice environment with the ability obtain immediate feedback for skill development. In addition to training, this device can be used in the future to assess new orthopaedic product designs, such as drills and guide fixtures. This expands the usability to a much larger application field. This research identifies specific training challenges for novices in the component skills of wire navigation, such as the estimating distances on radiographic images and assessing the anteversion of the femoral neck. This understanding presents new opportunities for research in training this critical surgical skill.

An important goal for future research on this topic is to find definitive evidence that training on the simulator transfers to a more realistic scenario. Due to residency size limitations and time required to acquire wire navigation skill, this is nearly impossible to complete in a single-day experiment. To complete this study, a standalone simulator needs to be created and implemented at an orthopaedic training institution for extended practice and further observation of a subject's ability to transfer simulated exercises to realistic surgical scenarios. Although more studies are needed in the future, this work makes substantial advancements in the orthopaedic field by developing a new and effective simulator for teaching surgical skill.

REFERENCES

[1] J. M. Froelich, J. C. Milbrandt, W. M. Novicoff, K. J. Saleh, and D. G. Allan, "Surgical simulators and hip fractures: a role in residency training?," *Journal of Surgical Education,* vol. 68, pp. 298-302, 2011.

[2] M. Vankipuram, K. Kahol, A. McLaren, and S. Panchanathan, "A virtual reality simulator for orthopedic basic skills: A design and validation study," *Journal of biomedical informatics,* vol. 43, pp. 661-668, 2010.

[3] T. A. Brennan, L. L. Leape, N. M. Laird, L. Hebert, A. R. Localio, A. G. Lawthers*, et al.*, "Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I," *New England journal of medicine,* vol. 324, pp. 370-376, 1991.

[4] (2010, 6/17/2013). *Deaths and Mortality - Centers for Disease Control and Prevention* [Online]. Available: http://www.cdc.gov/nchs/fastats/deaths.htm

[5] L. L. Leape, T. A. Brennan, N. Laird, A. G. Lawthers, A. R. Localio, B. A. Barnes*, et al.*, "The Nature of Adverse Events in Hospitalized Patients," *New England Journal of Medicine,* vol. 324, pp. 377-384, 1991.

[6] E. A. Halm, C. Lee, and M. R. Chassin, "Is volume related to outcome in health care? A systematic review and methodologic critique of the literature," *Annals of internal medicine,* vol. 137, pp. 511-520, 2002.

[7] R. H. Bell Jr, T. W. Biester, A. Tabuenca, R. S. Rhodes, J. B. Cofer, L. D. Britt*, et al.*, "Operative experience of residents in US general surgery programs: a gap between expectation and experience," *Annals of surgery,* vol. 249, pp. 719-724, 2009.

[8] J. L. Kelsey and S. Hoffman, "Risk factors for hip fracture," *The New England journal of medicine,* 1987.

[9] L. Melton III, S. Gabriel, C. Crowson, A. Tosteson, O. Johnell, and J. Kanis, "Cost-equivalence of different osteoporotic fractures," *Osteoporosis International,* vol. 14, pp. 383-388, 2003.

[10] D. M. Gaba, "The future vision of simulation in health care," *Quality and Safety in Health Care,* vol. 13, pp. i2-i10, 2004.

[11] A. Ziv, P. R. Wolpe, S. D. Small, and S. Glick, "Simulation-Based Medical Education: An Ethical Imperative," *Academic Medicine,* vol. 78, pp. 783-788, 2003.

[12] H. Owen, "Early use of simulation in medical education," *Simulation in Healthcare,* vol. 7, p. 102, 2012.

[13] M. A. Faria Jr, "The death of Henry II of France," *Journal of Neurosurgery,* vol. 77, pp. 964-969, 1992.

[14] "Advertisements," *New York medical journal* vol. 12, 1870.

[15] B. Howard, "Description of an apparatus for teaching hernia: its anatomy and mechanism," *New York Medical Journal,* vol. 8, pp. 46-55, 1868.

[16] S. W. Winchell and P. Safar, "Teaching and testing lay and paramedical personnel in cardiopulmonary resuscitation," *Anesthesia & Analgesia,* vol. 45, pp. 441-449, 1966.

[17] P. Safar, L. A. Escarraga, and J. O. Elam, "A Comparison of the Mouth-to-Mouth and Mouth-to-Airway Methods of Artificial Respiration with the Chest-Pressure Arm-Lift Methods," *New England Journal of Medicine,* vol. 258, pp. 671-677, 1958.

[18] J. Cooper and V. Taqueti, "A brief history of the development of mannequin simulators for clinical education and training," *Postgraduate medical journal,* vol. 84, pp. 563-570, 2008.

[19] M. S. Gordon, G. A. Ewy, A. C. DeLeon Jr, R. A. Waugh, J. M. Felner, A. D. Forker*, et al.*, "“Harvey,” the cardiology patient simulator: Pilot studies on teaching effectiveness," *The American Journal of Cardiology,* vol. 45, pp. 791-796, 1980.

[20] S. B. Issenberg, W. C. McGaghie, I. R. Hart, J. W. Mayer, J. M. Felner, E. R. Petrusa*, et al.*, "Simulation technology for health care professional skills training and assessment," *JAMA: the journal of the American Medical Association,* vol. 282, pp. 861-866, 1999.

[21] M. Wilson, A. Middlebrook, C. Sutton, R. Stone, and R. McCloy, "MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance," *Annals of the Royal College of Surgeons of England,* vol. 79, p. 403, 1997.

[22] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. O’Brien, V. K. Bansal, D. K. Andersen*, et al.*, "Virtual reality training improves operating room performance: results of a randomized, double-blinded study," *Annals of surgery,* vol. 236, p. 458, 2002.

[23] S. N. Kothari, B. J. Kaplan, E. J. DeMaria, T. J. Broderick, and R. C. Merrell, "Training in laparoscopic suturing skills using a new computer-based virtual reality simulator (MIST-VR) provides results comparable to those with an established pelvic trainer system," *Journal of Laparoendoscopic & Advanced Surgical Techniques,* vol. 12, pp. 167-173, 2002.

[24] A. G. Gallagher, K. Richie, N. McClure, and J. McGuigan, "Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality," *World journal of surgery,* vol. 25, pp. 1478-1483, 2001.

[25] J. Torkington, S. Smith, B. Rees, and A. Darzi, "Skill transfer from virtual reality to a real laparoscopic task," *Surgical endoscopy,* vol. 15, pp. 1076-1079, 2001.

[26] A. Chaudhry, C. Sutton, J. Wood, R. Stone, and R. McCloy, "Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface," *Annals of the Royal College of Surgeons of England,* vol. 81, p. 281, 1999.

[27] A. M. Derossis, G. M. Fried, M. Abrahamowicz, H. H. Sigman, J. S. Barkun, and J. L. Meakins, "Development of a model for training and evaluation of laparoscopic skills," *The American journal of surgery,* vol. 175, pp. 482-487, 1998.

[28] M. Vassiliou, G. Ghitulescu, L. Feldman, D. Stanbridge, K. Leffondre, H. Sigman*, et al.*, "The MISTELS program to measure technical skill in laparoscopic surgery," *Surgical Endoscopy And Other Interventional Techniques,* vol. 20, pp. 744-747, 2006.

[29] G. Fried, A. Derossis, J. Bothwell, and H. Sigman, "Comparison of laparoscopic performance in vivo with performance measured in a laparoscopic simulator," *Surgical endoscopy,* vol. 13, pp. 1077-1081, 1999.

[30] J. Peters, G. M. Fried, L. L. Swanstrom, N. J. Soper, L. F. Sillin, B. Schirmer*, et al.*, "Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery," *Surgery,* vol. 135, pp. 21-27, 2004.

[31] K. Kunkler, "The role of medical simulation: an overview," *The International Journal of Medical Robotics and Computer Assisted Surgery,* vol. 2, pp. 203-210, 2006.

[32] R. H. Bell Jr, "Why Johnny cannot operate," *Surgery,* vol. 146, pp. 533-542, 2009.

[33] C. B. Franzese and S. P. Stringer, "The evolution of surgical training: perspectives on educational models from the past to the future," *Otolaryngologic Clinics of North America,* vol. 40, pp. 1227-1235, 2007.

[34] J. L. Cameron, "William Stewart Halsted. Our surgical heritage," *Annals of Surgery,* vol. 225, p. 445, 1997.

[35] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The role of deliberate practice in the acquisition of expert performance," *Psychological review,* vol. 100, p. 363, 1993.

[36] F. Galton, *Hereditary genius*: Macmillan and Company, 1869.

[37] E. E. Ghiselli, *The validity of occupational aptitude tests* vol. 8: Wiley, 1966.

[38] E. E. Ghiselli, "The Validity of Aptitude Tests in Personnel Selection," *Personnel Psychology,* vol. 26, pp. 461-477, 1973.

[39] K. R. Wanzel, S. J. Hamstra, D. J. Anastakis, E. D. Matsumoto, and M. D. Cusimano, "Effect of visual-spatial ability on learning of spatially-complex surgical skills," *The lancet,* vol. 359, pp. 230-231, 2002.

[40] J. Y. Lee, D. C. Kerbl, E. M. McDougall, and P. Mucksavage, "Medical Students Pursuing Surgical Fields Have No Greater Innate Motor Dexterity than Those Pursuing Nonsurgical Fields," *Journal of Surgical Education,* vol. 69, pp. 360-363, 2012.

[41] K. R. Wanzel, S. J. Hamstra, M. F. Caminiti, D. J. Anastakis, E. D. Grober, and R. K. Reznick, "Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance," *Surgery,* vol. 134, pp. 750-757, 2003.

[42] C. J. Harris, M. Herbert, and R. J. C. Steele, "Psychomotor skills of surgical trainees compared with those of different medical specialists," *British Journal of Surgery,* vol. 81, pp. 382-383, 1994.

[43] N. K. Francis, G. B. Hanna, A. B. Cresswell, F. J. Carter, and A. Cuschieri, "The performance of master surgeons on standard aptitude testing," *The American journal of surgery,* vol. 182, pp. 30-33, 2001.

[44] L. L. Baird, "Do grades and tests predict adult accomplishment?," *Research in Higher Education,* vol. 23, pp. 3-85, 1985.

[45] K. A. Ericsson and N. Charness, "Expert performance," *American psychologist,* vol. 49, pp. 725-747, 1994.

[46] H. S. Luft, J. P. Bunker, and A. C. Enthoven, "Should operations be regionalized? the empirical relation between surgical volume and mortality," *The New England Journal of Medicine,* vol. 301, p. 1364, 1979.

[47] H. S. Luft, "The relation between surgical volume and mortality: an exploration of causal factors and alternative models," *Medical Care,* pp. 940-959, 1980.

[48] H. S. Luft, S. S. Hunt, and S. C. Maerki, "The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?," *Health services research,* vol. 22, p. 157, 1987.

[49] A. B. Flood, W. R. Scott, and W. Ewy, "Does practice make perfect?: Part I: the relation between hospital volume and outcomes for selected diagnostic categories," *Medical care,* vol. 22, pp. 98-114, 1984.

[50] A. B. Flood, W. R. Scott, and W. Ewy, "Does Practice Make Perfect? Part II: The Relation between Volume and and Outcomes and Other Hospital Characteristics," *Medical Care,* pp. 115-125, 1984.

[51] J. D. Birkmeyer, A. E. Siewers, E. V. A. Finlayson, T. A. Stukel, F. L. Lucas, I. Batista*, et al.*, "Hospital Volume and Surgical Mortality in the United States," *New England Journal of Medicine,* vol. 346, pp. 1128-1137, 2002.

[52] A. J. Vickers, F. J. Bianco, A. M. Serio, J. A. Eastham, D. Schrag, E. A. Klein*, et al.*, "The surgical learning curve for prostate cancer control after radical prostatectomy," *Journal of the National Cancer Institute,* vol. 99, pp. 1171-1177, 2007.

[53] C. B. Begg, E. R. Riedel, P. B. Bach, M. W. Kattan, D. Schrag, J. L. Warren*, et al.*, "Variations in morbidity after radical prostatectomy," *New England Journal of Medicine,* vol. 346, pp. 1138-1144, 2002.

[54] K. Bjorgul, W. M. Novicoff, and K. J. Saleh, "Learning curves in hip fracture surgery," *International orthopaedics,* vol. 35, pp. 113-119, 2011.

[55] H. D. Taylor, D. A. Dennis, and H. S. Crane, "Relationship between mortality rates and hospital patient volume for Medicare patients undergoing major orthopaedic surgery of the hip, knee, spine, and femur," *The Journal of arthroplasty,* vol. 12, pp. 235-242, 1997.

[56] F. J. Bianco, A. M. Cronin, E. A. Klein, J. E. Pontes, P. T. Scardino, and A. J. Vickers, "Fellowship Training as a Modifier of the Surgical Learning Curve," *Academic Medicine,* vol. 85, pp. 863-868, 2010.

[57] F. S. Keller, "The Phantom Plateau," *Journal of the Experimental Analysis of Behavior,* vol. 1, pp. 1-13, 1958.

[58] (2013, 6/10/2013). *Top Finishers - Boston Athletic Association*. Available: http://raceday.baa.org/top-finishers.html

[59] H. A. Simon and W. G. Chase, "Skill in chess: Experiments with chess-playing tasks and computer simulation of skilled performance throw light on some human perceptual and memory processes," *American scientist,* vol. 61, pp. 394-403, 1973.

[60] N. V. Krogius, *Psychology in chess*: RHM Press, 1976.

[61] E. L. Thorndike, "Educational psychology, Vol 2: The psychology of learning," 1913.

[62] K. A. Ericsson, "The influence of experience and deliberate practice on the development of superior expert performance," *The Cambridge handbook of expertise and expert performance,* pp. 683-703, 2006.

[63] M. H. Trowbridge and H. Cason, "An experimental study of Thorndike's theory of learning," *The Journal of General Psychology,* vol. 7, pp. 245-260, 1932.

[64] J. S. Butterworth and E. H. Reppert, "Auscultatory acumen in the general medical population," *JAMA: The Journal of the American Medical Association,* vol. 174, pp. 32-34, 1960.

[65] C.-A. E. Moulton, A. Dubrowski, H. MacRae, B. Graham, E. Grober, and R. Reznick, "Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial," *Annals of surgery,* vol. 244, p. 400, 2006.

[66] F. N. Dempster, "Spacing effects and their implications for theory and practice," *Educational Psychology Review,* vol. 1, pp. 309-330, 1989.

[67] M. Pusic, M. Pecaric, and K. Boutis, "How much practice is enough? Using learning curves to assess the deliberate practice of radiograph interpretation," *Academic Medicine,* vol. 86, p. 731, 2011.

[68] K. A. Ericsson, "Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains," *Academic Medicine,* vol. 79, pp. S70-S81, 2004.

[69] W. C. McGaghie, S. B. Issenberg, M. E. R. Cohen, J. H. Barsuk, and D. B. Wayne, "Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence," *Academic medicine: journal of the Association of American Medical Colleges,* vol. 86, p. 706, 2011.

[70] W. C. McGaghie, V. J. Siddall, P. E. Mazmanian, and J. Myers, "Lessons for Continuing Medical Education From Simulation Research in Undergraduate and Graduate Medical EducationEffectiveness of Continuing Medical Education: American College of Chest Physicians Evidence-Based Educational Guidelines," *Chest Journal,* vol. 135, pp. 62S-68S, 2009.

[71] D. M. Berwick and L. L. Leape, "Reducing errors in medicine: it's time to take this more seriously," *BMJ: British Medical Journal,* vol. 319, p. 136, 1999.

[72] H. H. McGuire Jr, J. S. Horsley III, D. R. Salter, and M. Sobel, "Measuring and managing quality of surgery: statistical vs incidental approaches," *Archives of Surgery,* vol. 127, p. 733, 1992.

[73] E. J. Thomas, D. M. Studdert, H. R. Burstin, E. J. Orav, T. Zeena, E. J. Williams*, et al.*, "Incidence and types of adverse events and negligent care in Utah and Colorado," *Medical care,* vol. 38, pp. 261-271, 2000.

[74] T. A. Brennan, "Analysis of errors reported by surgeons at three teaching hospitals," *Surgery,* vol. 133, pp. 614-21, 2003.

[75] L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, *To err is human: building a safer health system* vol. 627: National Academies Press, 2000.

[76] D. V. Gelfand, Y. D. Podnos, J. C. Carmichael, D. J. Saltzman, S. E. Wilson, and R. A. Williams, "Effect of the 80-hour workweek on resident burnout," *Archives of Surgery,* vol. 139, p. 933, 2004.

[77] B. Weatherby, J. Rudd, T. Ervin, P. Stafford, and B. Norris, "The effect of resident work hour regulations on orthopaedic surgical education," *Journal of surgical orthopaedic advances,* vol. 16, p. 19, 2007.

[78] M. M. Hutter, K. C. Kellogg, C. M. Ferguson, W. M. Abbott, and A. L. Warshaw, "The impact of the 80-hour resident workweek on surgical residents and attending surgeons," *Annals of surgery,* vol. 243, p. 864, 2006.

[79] C. P. Landrigan, L. K. Barger, B. E. Cade, N. T. Ayas, and C. A. Czeisler, "Interns' compliance with accreditation council for graduate medical education work-hour limits," *JAMA: the journal of the American Medical Association,* vol. 296, pp. 1063-1070, 2006.

[80] R. K. Chandra, "The Resident 80−Hour Work Week: How Has It Affected Surgical Specialties?," *The Laryngoscope,* vol. 114, pp. 1394-1398, 2004.

[81] K. A. Ericsson and A. C. Lehmann, "Expert and exceptional performance: Evidence of maximal adaptation to task constraints," *Annual review of psychology,* vol. 47, pp. 273-305, 1996.

[82] R. S. Chung, "How much time do surgical residents need to learn operative surgery?," *The American journal of surgery,* vol. 190, pp. 351-353, 2005.

[83] M. Bridges and D. L. Diamond, "The financial impact of teaching surgical residents in the operating room," *The American Journal of Surgery,* vol. 177, pp. 28-32, 1999.

[84] R. D. Shìppert, "A study of time-dependent operating room fees and how to save $100,000 by using time-saving products," *Am J Cosmet Surg,* vol. 22, pp. 25-34, 2005.

[85] T. J. Babineau, J. Becker, G. Gibbons, S. Sentovich, D. Hess, S. Robertson*, et al.*, "The "cost" of operative training for surgical residents," *Archives of Surgery,* vol. 139, p. 366, 2004.

[86] L. R. Farnworth, D. E. Lemay, T. Wooldridge, J. D. Mabrey, M. Blaschak, T. A. DeCoster*, et al.*, "A comparison of operative times in arthroscopic ACL reconstruction between orthopaedic faculty and residents: the financial impact of orthopaedic surgical training in the operating room," *The Iowa orthopaedic journal,* vol. 21, p. 31, 2001.

[87] P. B. Andreatta, D. T. Woodrum, J. D. Birkmeyer, R. K. Yellamanchilli, G. M. Doherty, P. G. Gauger*, et al.*, "Laparoscopic skills are improved with LapMentor™ training: results of a randomized, double-blinded study," *Annals of surgery,* vol. 243, p. 854, 2006.

[88] K. A. Ericsson, "An expert−performance perspective of research on medical expertise: the study of clinical performance," *Medical Education,* vol. 41, pp. 1124-1130, 2007.

[89] *Standards for Educational & Psychological Tests*: American Psychological Association, American Educational Research Association. National Council on Measurement in Education, 1974.

[90] T. D. Cook, D. T. Campbell, and A. Day, *Quasi-experimentation: Design & analysis issues for field settings*: Houghton Mifflin Boston, 1979.

[91] *Standards for educational and psychological testing*: American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational Psychological Testing, 1999.

[92] F. Carter, M. Schijven, R. Aggarwal, T. Grantcharov, N. Francis, G. Hanna*, et al.*, "Consensus guidelines for validation of virtual reality surgical simulators," *Surgical Endoscopy and Other Interventional Techniques,* vol. 19, pp. 1523-1532, 2005.

[93] B. Dauster, A. P. Steinberg, M. C. Vassiliou, S. Bergman, D. D. Stanbridge, L. S. Feldman*, et al.*, "Validity of the MISTELS simulator for laparoscopy training in urology," *Journal of endourology,* vol. 19, pp. 541-545, 2005.

[94] A. Derossis, J. Bothwell, H. Sigman, and G. Fried, "The effect of practice on performance in a laparoscopic simulator," *Surgical endoscopy,* vol. 12, pp. 1117-1120, 1998.

[95] A. Sourin, O. Sourina, and H. T. Sen, "Virtual orthopedic surgery training," *Computer Graphics and Applications, IEEE,* vol. 20, pp. 6-9, 2000.

[96] O. Sourina, A. Sourin, and H. T. Sen, "Virtual orthopedic surgery training on personal computer," *International Journal of Information Technology,* vol. 6, pp. 16-29, 2000.

[97] M.-D. Tsai, M.-S. Hsieh, and S.-B. Jou, "Virtual reality orthopedic surgery simulator," *Computers in biology and medicine,* vol. 31, pp. 333-351, 2001.

[98] P. Blyth, N. Stott, and I. Anderson, "A simulation-based training system for hip fracture fixation for use within the hospital environment," *Injury,* vol. 38, pp. 1197-1203, 2007.

[99] P. Blyth, N. Stott, and I. Anderson, "Virtual reality assessment of technical skill using the Bonedoc DHS simulator," *Injury,* vol. 39, pp. 1127-1133, 2008.

[100] M.-D. Tsai, M.-S. Hsieh, and C.-H. Tsai, "Bone drilling haptic interaction for orthopedic surgical simulator," *Computers in Biology and Medicine,* vol. 37, pp. 1709-1718, 2007.

[101] J. Pettersson, K. L. Palmerius, H. Knutsson, O. Wahlstrom, B. Tillander, and M. Borga, "Simulation of Patient Specific Cervical Hip Fracture Surgery With a Volume Haptic Interface," *Biomedical Engineering, IEEE Transactions on,* vol. 55, pp. 1255-1265, 2008.

[102] P. Blyth, I. A. Anderson, and N. S. Stott, "Virtual reality simulators in orthopedic surgery: what do the surgeons think?," *Journal of Surgical Research,* vol. 131, pp. 133-139, 2006.

[103] G. Thomas, L. Johnson, S. Dow, and C. Stanford, "The design and testing of a force feedback dental simulator," *Computer methods and programs in biomedicine,* vol. 64, pp. 53-64, 2001.

[104] S. Dow, G. Thomas, and L. Johnson, "Signal detection performance with a haptic device," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1999, pp. 1233-1237.

[105] G. W. Thomas and L. Liu, "Displaying Small Surface Features with a Force Feedback Device in a Dental Training Simulator," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2005, pp. 2235-2239.

[106] J. E. Colgate and J. M. Brown, "Factors affecting the z-width of a haptic display," in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, 1994, pp. 3205-3210.

[107] J. E. Colgate, M. C. Stanley, and J. M. Brown, "Issues in the haptic display of tool use," in *Human Robot Interaction and Cooperative Robots, Proceedings. 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems.*, 1995, pp. 140-145.

[108] W. R. Mark, S. C. Randolph, M. Finch, J. M. Van Verth, and R. M. Taylor II, "Adding force feedback to graphics systems: Issues and solutions," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 1996, pp. 447-452.

[109] D. Morris, C. Sewell, N. Blevins, F. Barbagli, and K. Salisbury, "A collaborative virtual environment for the simulation of temporal bone surgery," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, ed: Springer, 2004, pp. 319-327.

[110] R. Marks, J. P. Allegrante, C. Ronald MacKenzie, and J. M. Lane, "Hip fractures among the elderly: causes, consequences and control," *Ageing research reviews,* vol. 2, pp. 57-93, 2003.

[111] S. R. Cummings, S. M. Rubin, and D. Black, "The future of hip fractures in the United States: numbers, costs, and potential effects of postmenopausal estrogen," *Clinical Orthopaedics and Related Research,* vol. 252, pp. 163-166, 1990.

[112] A. N. Tosteson, R. T. Burge, D. A. Marshall, and R. Lindsay, "Therapies for Treatment of Osteoporosis in US Women: Cost-effectiveness and Budget Impact Considerations-Page 4," *Am J Manag Care,* vol. 14, pp. 605-615, 2008.

[113] C. Cooper, G. Campion, and L. Melton III, "Hip fractures in the elderly: a world-wide projection," *Osteoporosis international,* vol. 2, pp. 285-289, 1992.

[114] W. E. Garrett Jr, M. F. Swiontkowski, J. N. Weinstein, J. Callaghan, R. N. Rosier, D. J. Berry*, et al.*, "American Board of Orthopaedic Surgery Practice of the Orthopaedic Surgeon: Part-II, Certification Examination Case Mix," *The Journal of Bone & Joint Surgery,* vol. 88, pp. 660-667, 2006.

[115] G. Bannister, A. Gibson, C. Ackroyd, and J. Newman, "The fixation and prognosis of trochanteric fractures: a randomized prospective controlled trial," *Clinical orthopaedics and related research,* vol. 254, pp. 242-246, 1990.

[116] G. L. Wolfgang, M. H. Bryant, and J. P. O'Neill, "Treatment of intertrochanteric fracture of the femur using sliding screw plate fixation," *Clinical orthopaedics and related research,* vol. 163, pp. 148-158, 1982.

[117] S. R. Cummings, D. M. Black, and S. M. Rubin, "Lifetime risks of hip, Colles', or vertebral fracture and coronary heart disease among white postmenopausal women," *Archives of Internal Medicine,* vol. 149, p. 2445, 1989.

[118] C. A. Brauer, M. Coca-Perraillon, D. M. Cutler, and A. B. Rosen, "Incidence and mortality of hip fractures in the United States," *JAMA: the journal of the American Medical Association,* vol. 302, pp. 1573-1579, 2009.

[119] E. J. Samelson, Y. Zhang, D. P. Kiel, M. T. Hannan, and D. T. Felson, "Effect of birth cohort on risk of hip fracture: age-specific incidence rates in the Framingham Study," *American Journal of Public Health,* vol. 92, pp. 858-862, 2002.

[120] R. Garden, "Low-angle fixation in fractures of the femoral neck," *Journal of Bone & Joint Surgery, British Volume,* vol. 43, pp. 647-663, 1961.

[121] E. M. Evans, "The treatment of trochanteric fractures of the femur," *Journal of Bone & Joint Surgery, British Volume,* vol. 31, pp. 190-203, 1949.

[122] J. S. Jensen and M. Michaelsen, "Trochanteric femoral fractures treated with McLaughlin osteosynthesis," *Acta Orthopaedica,* vol. 46, pp. 795-803, 1975.

[123] *Hip Fractures - OrthoInfo - AAOS*. Available: http://orthoinfo.aaos.org/topic.cfm?topic=A00392

[124] M. Butler, M. Forte, R. L. Kane, S. Joglekar, S. J. Duval, M. Swiontkowski*, et al.*, *Treatment of common hip fractures*: Agency for Healthcare Research and Quality, 2009.

[125] K. Ohzono, M. Saito, K. Takaoka, K. Ono, S. Saito, T. Nishina*, et al.*, "Natural history of nontraumatic avascular necrosis of the femoral head," *Journal of Bone & Joint Surgery, British Volume,* vol. 73-B, pp. 68-72, January 1, 1991 1991.

[126] M. Ho, G. Garau, G. Walley, F. Oliva, A. S. Panni, U. G. Longo*, et al.*, "Minimally invasive dynamic hip screw for fixation of hip fractures," *International orthopaedics,* vol. 33, pp. 555-560, 2009.

[127] N. Suhm, P. Messmer, I. Zuna, L. A. Jacob, and P. Regazzoni, "Fluoroscopic guidance versus surgical navigation for distal locking of intramedullary implants: A prospective, controlled clinical study," *Injury,* vol. 35, pp. 567-574, 2004.

[128] R. Sanders, K. Koval, T. DiPasquale, G. Schmelling, S. Stenzler, and E. Ross, "Exposure of the orthopaedic surgeon to radiation," *Journal of Bone & Joint Surgery, American Volume,* vol. 75, pp. 326-326, 1993.

[129] J. Coetzee and E. Van Der Merwe, "Exposure of surgeons-in-training to radiation during intramedullary fixation of femoral shaft fractures," *South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde,* vol. 81, p. 312, 1992.

[130] M. Noorden, C. Lavy, T. Briggs, and M. Roos, "Unrecognised joint penetration in treatment of femoral neck fractures," *Journal of Bone & Joint Surgery, British Volume,* vol. 75, pp. 448-448, 1993.

[131] J. Bahner-Heyne, S. Roettger, D. Schulze-Kissing, and D. Manzey, "Automation in Surgery: The Surgeons' Perspective on Human Factors Issues of Image-Guided Navigation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2009, pp. 679-683.

[132] J. P. Cobb, V. Kannan, K. Brust, and G. Thevendran, "Navigation reduces the learning curve in resurfacing total hip arthroplasty," *Clinical orthopaedics and related research,* vol. 463, pp. 90-97, 2007.

[133] M. Liebergall, R. Mosheiff, and D. Segal, "Navigation in orthopaedic trauma," *Operative Techniques in Orthopaedics,* vol. 13, pp. 64-72, 2003.

[134] D. Manzey, S. Röttger, J. E. Bahner−Heyne, D. Schulze−Kissing, A. Dietz, J. Meixensberger*, et al.*, "Image−guided navigation: the surgeon's perspective on performance consequences and human factors issues," *The International Journal of Medical Robotics and Computer Assisted Surgery,* vol. 5, pp. 297-308, 2009.

[135] M. C. Müller, P. Belei, P. H. Pennekamp, K. Kabir, D. C. Wirtz, C. Burger*, et al.*, "Three-dimensional computer-assisted navigation for the placement of cannulated hip screws. A pilot study," *International orthopaedics,* vol. 36, pp. 1463-1469, 2012.

[136] R. Rambani, W. Viant, J. Ward, and A. Mohsen, "Computer-Assisted Orthopedic Training System for Fracture Fixation," *Journal of Surgical Education,* 2013.

[137] E. Sidon and E. L. Steinberg, "Accuracy study of new computer-assisted orthopedic surgery software," *European Journal of Radiology,* 2012.

[138] R. Gundle, M. Gargan, and A. Simpson, "How to minimize failures of fixation of unstable intertrochanteric fractures," *Injury,* vol. 26, pp. 611-614, 1995.

[139] S. Larsson, S. Friberg, and L.-I. Hansson, "Trochanteric fractures: influence of reduction and implant position on impaction and complications," *Clinical orthopaedics and related research,* vol. 259, pp. 130-139, 1990.

[140] G. S. Laros and J. F. Moore, "Complications of fixation in intertrochanteric fractures," *Clinical orthopaedics and related research,* vol. 101, pp. 110-119, 1974.

[141] H. Kaufer, "Mechanics of the treatment of hip injuries," *Clinical Orthopaedics and Related Research,* vol. 146, pp. 53-61, 1980.

[142] M. Baumgaertner, S. Curtin, D. Lindskog, and J. Keggi, "The value of the tip-apex distance in predicting failure of fixation of the hip," *J Bone Joint Surg Am,* vol. 77, pp. 1058-1064, 1995.

[143] M. R. Baumgaertner and B. D. Solberg, "Awareness of tip-apex distance reduces failure of fixation of trochanteric fractures of the hip," *Journal of Bone & Joint Surgery, British Volume,* vol. 79, pp. 969-971, 1997.

[144] J. A. Geller, C. Saifi, T. A. Morrison, and W. Macaulay, "Tip-apex distance of intramedullary devices as a predictor of cut-out failure in the treatment of peritrochanteric elderly hip fractures," *International orthopaedics,* vol. 34, pp. 719-722, 2010.

[145] J. Caird, "Persistent issues in the application of virtual environment systems to training," in *Human Interaction with Complex Systems, 1996. HICS'96. Proceedings., Third Annual Symposium on*, 1996, pp. 124-132.

[146] G. Lintern, S. N. Roscoe, and J. E. Sivier, "Display principles, control dynamics, and environmental factors in pilot training and transfer," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 32, pp. 299-317, 1990.

[147] G. Lintern, S. N. Roscoe, J. M. Koonce, and L. D. Segal, "Transfer of landing skills in beginning flight training," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 32, pp. 319-327, 1990.

[148] G. Lintern, "An informational perspective on skill transfer in human-machine systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 33, pp. 251-266, 1991.

[149] J. M. Kracht, I. J. Busch-Vishniac, and J. E. West, "Noise in the operating rooms of Johns Hopkins Hospital," *The Journal of the Acoustical Society of America,* vol. 121, pp. 2673-2680, 2007.

[150] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: a class of displays on the reality-virtuality continuum," pp. 282-292, 1995.

[151] G. E. Stelmach and A. Sirica, "Aging and proprioception," *Age,* vol. 9, pp. 99-103, 1986.

[152] R. J. van Beers, A. C. Sittig, and J. J. van der Gon Denier, "How humans combine simultaneous proprioceptive and visual position information," *Experimental Brain Research,* vol. 111, pp. 253-261, 1996.

[153] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology,* vol. 47, p. 381, 1954.

[154] O. Sourina, A. Sourin, and H. T. Sen, "Orthopedic Surgery Training Simulation," *Journal of Mechanics in Medicine and Biology,* vol. 7, pp. 37-53, 2007.

[155] P. Blyth and P. Sehgal, "Use of the Bonedoc DHS simulator by fifth year medical students: A pilot study," in *ASCILITE-Australian Society for Computers in Learning in Tertiary Education Annual Conference*, 2009, pp. 74-80.