An Objective Methodology to Quantify Motor Skills in Basic Orthopaedic and Gynecologic Surgical Tasks

A Thesis submitted to The University of Arizona College of Medicine-Phoenix in partial fulfillment of the requirements for the Degree of Doctor of Medicine.

Sarah Kennedy Sharer

2011

Dedication

This work is dedicated to the faculty and administration of the University of Arizona College of Medicine—Phoenix; especially those who envisioned the scholarly work of the inaugural class, long before our presence on campus.

Acknowledgements

The author would like to thank Kanav Kahol Ph.D. for his time and commitment to this project, without all of his help this project would not have been possible. Also, the author would like to thank Matthew Nugent M.D., Alex McLaren M.D., Melissa Gioia M.D., Mithra Vankipuram M.S. and Foad Saeidi for their contributions to the projects. Funding and technical help were secured through a Banner Health research grant and Arizona State University. Equipment was provided by the SimET Center at Banner Good Samaritan Medical Center and Synthes®. Anatomic specimens were provided by Banner Good Samaritan Medical Center.

Abstract

This study is the initial step in the development of a basic skills simulator for surgical residents, intended to serve as a cross-platform objective motor skills training system. A robust system was developed to measure motor skills of orthopaedic surgeons and gynecologic surgeons. The orthopaedic study focused on three basic skills: drill, tap, and screw insertion. Each of the participants repeated the sequential task of drill, tap, screw insertion ten times in cadaveric femoral diaphyseal bone. The gynecologic study focused on placement of sutures across ten incisions in synthetic skin. Real time wrist, hand, and finger position was recorded bilaterally using Immersion CyberGloves® and the Ascension Liberty Tracker®. Four metrics were evaluated: task duration, gesture proficiency, hand movement smoothness (replaced by subjective rating in Gyn study), and number of errors, in order to differentiate between groups with different skill levels: No Skills, Limited Skills, Advanced Skills, and Expert Skills. The results demonstrated a significant difference between the Expert Skills group and the less skilled groups. In both studies the Expert Skills group completed the tasks the fastest, with the greatest proficiency, smoothness and accuracy. These results establish the

capability of our present skill quantification system and will assist in future improvement of the analysis algorithm.

Table of Contents

List of Figures and Tables

Introduction

Virtual reality simulation is receiving increasing attention as a tool for use in physician training [1,2]. With a decrease in resident work hours and an increased focus on patient safety and cost reduction, new methods of resident training are being encouraged [2-4]. Simulation training has been introduced for technical procedures in General Surgery, Obstetrics, and Anesthesia through simulators for laparoscopy, birthing, and fiberoptic intubation, respectively [5-9]. Orthopaedics, however, is behind in its use of computer based simulators in training curricula. The majority of simulator training in orthopaedics has been focused on arthroscopic surgery. Arthroscopic simulators generally encompass entire procedures rather than focusing on basic motor skills such as drilling or suturing and there is a lack of trainers for other basic orthopaedic skills. Unlike procedural simulators that aim to simulate an entire procedure, basic skills trainers focus on individual skills (drilling, sawing, triangulation), providing users with practice exercises on these basic skills. In most cases, both procedural and basic skills trainers can contribute effectively to resident training. An example of a basic skills trainer is the Fundamentals of Laparoscopy Trainer Box (SAGES/FLS, Los

Angeles, CA) which is a validated trainer in general surgery [10]. The American Board of Surgery now requires each resident to clear skills exams on this trainer to become a board certified surgeon [11]. Obstetrics and Gynecology residency programs are now utilizing the FLS trainer for resident basic skill (laparoscopy) practice for gynecologic surgeries.

Skill level of surgeons is an important metric to capture in a training environment. To date, teaching of skills has been subjective. However, with the aid of sophisticated tools and algorithms there is a possibility of quantifying surgical skills, identifying errors and providing formative and summative feedback to the residents on their skill levels. This is very useful in a training environment and enables users to improve their skills in a fast and efficient manner.

Performance outcome measures are used to evaluate the fine motor skills displayed during the completion of surgical tasks. In order to objectively assess performance, appropriate metrics must be defined. According to Gallagher, metrics formulation requires task deconstruction and subsequent definition of optimal and suboptimal performance [12]. Objective outcome measures in surgery are identified in the literature as the Objective Structured Assessment of

Technical Skills (OSATS), and use of the Imperial College Surgical Assessment Device (ICSAD) [13-15]. Both are considered effective in differentiating levels of expertise in surgical tasks [13]. An experienced surgeon demonstrates his or her expertise through sophisticated motor skills. An expert's performance appears refined, smooth, and very efficient. Comparatively, the novice's performance in the same situation appears awkward and highly inefficient [16]. Datta's study of manual dexterity and outcome measures, evaluated using the electromagnetic tracking system of the ICSAD, suggested that the outcome of a general surgery procedure can be predicted by measuring surgical skill [17]. Although these methods have not been employed for gynecologic or orthopaedic skill quantification, it is likely that such techniques could readily be used in gynecologic skill assessment (due to the similarity in skills), whereas the transfer to orthopaedics would be less plausible.

When developing a skills quantification system for orthopaedics, certain factors need to be accounted for. First, orthopaedic skills are considerably different than general surgery skills in terms of movement scope, range of forces and intended effect. While surgical skills trainers have been developed and used to evaluate performance and skill level, those trainers do not simulate the skills used in orthopaedic surgery.   Secondly it is important to, at-least in theory, build a system that can not only quantify skills in simulation environments but also in actual surgical environments. In general there is a lack of well defined objective measures for surgical proficiency that enable standardized testing in both virtual reality (VR) and the operating room (OR). This factor severely inhibits conducting control studies, which objectively establish learning curves and evaluations. In a simple design, senior surgeons could be requested to review resident's performance in simulation task and use the same scale for performance evaluation in real surgery like an extension of the aforementioned OSATS system. Experience in this domain however shows that often real time analysis of skill is not sufficient, and sessions need to be recorded and analyzed offline for complete evaluations and to provide feedback to residents on their performance. However, skill transfer and generalization from simulation to actual surgery are difficult to measure with such methodology due to reluctance of hospitals and patients to videotape actual surgery. Evaluation technologies currently employed in simulators, including the arthroscopy simulators, are not utilized in analysis of skills in the operating room because the technology is not easily transferred into the OR setting. This leaves many studies with the option of coarse level analysis of proficiency in actual procedures and correlating them to performance in surgical simulation training. It must be noted however that fine level detailed analysis in actual surgical procedures is necessary to reveal important mechanisms involved in surgical learning. In general, there is a need to design, develop and evaluate a cross compatible evaluation system that can be employed with many simulators and many skills and even in actual surgery.

When developing a skill evaluation system, behaviorally we could choose to (a) instrument the tools or (b) instrument the hand. In most simulators, designers instrument the tools. This allows the simulation to capture movements of the tools and analyze those to determine surgical proficiency. However instrumented tools cannot be employed in actual surgery and further cannot be employed across simulators. An alternative approach would be to instrument the hand. A surgeon's hand movements and shoulder movements could be captured through movement sensors and be potentially employed for quantification of skills. This system would focus on the surgeon's movements and not on the tool movement. It would allow for a cross-platform objective proficiency measurement system. While concerns on patient safety, infection and sterility need to be addressed when considering using such a system in surgery, if properly addressed such a system could be a useful resource in training of residents. This proficiency measurement system works with data gloves that can capture hand movements of surgeons to estimate skill. These gloves are unobtrusive systems for measuring proficiency with any simulator or surgical procedure. This system will permit detailed capture of surgical procedures and allow analysis and feedback by the surgeons.

A literature search revealed no data delineating hand dexterity during basic orthopaedic or gynecologic skills or correlation of hand dexterity with skill level. Studies conducted by Tashiro, Woodrow, Leong, and Howells [11,18-20] (as further explained in the discussion) employ force profiles and tool movements as means of assessing competency of the performer in simulation environments and some type of observations in the OR. However, these studies fail to measure hand movements and posture during the performance of the tasks. This failure proves to be a major limitation of these studies as it can be argued that hand movement and posture form the basis of successful performance of basic skills.

There are several advantages of measuring hand movements. Hand movements allow for detailed analysis and feedback. For instance, considering drilling, an algorithm that is based on assessing forces and tool motion can objectively analyze whether an operator performed the task satisfactorily. However it will only have limited feedback capacity as it may be able to only tell the user that the drilling angle was wrong or that too much or too little force was applied. It will not be able to assess whether the equipment was held in the hand with proper posture or whether unnecessary movements of the hands yielded inefficient movement. By using a data glove system it is possible to identify such factors in addition to all the factors associated with tool movements (which can be estimated by the hand movement profiles). Another situation in which measurement of hand movements yields greater information is in suturing. The precision that is required with suturing of skin, vessels, or ligaments is best determined through measurement of the angles and movements of the hands and wrists (and thus tool movements can be derived/estimated from the results).

The purpose of these studies is to capture motor skills data for drilling, thread tapping and screw insertion, and suturing utilizing the Immersion Cybergloves® and to correlate that data with performance level and learning of the task.

Aims

The ultimate aim of this study is to improve future utilization of surgical simulation training in residencies of various surgical disciplines. In order to accomplish this, better simulators must be developed. This work aims to develop a robust system to measure skill proficiency in the basic skills of surgery. In this study hand movement and posture data will be captured during three basic orthopaedic motor skills tasks and one basic surgical skill—suturing. The data will be then be correlated with skill proficiency for each level of training. Eventually the measurement of proficiency could then be incorporated into the design of a new simulator.

Hypothesis

Surgical skill levels of individuals with various years of training can be differentiated using data gloves (Immersion CyberGloves ®).

Research Materials and Methods

Study 1

Institutional Review Board approval was obtained through Banner Health and informed consent was obtained from all participants. Three groups of participants were recruited based on experience:

1- No Skills (medical students that had never used a drill);

2- Limited Skills (junior orthopaedic residents with six months to one year of exposure to these skills);

3- Expert Skills (practicing orthopaedic surgeons with extensive use of screws as a routine part of their practice).

A total of sixteen participants were enrolled; six in the No Skills group, four in the Limited Skills group, and six in the Expert Skills group. Power analysis was performed to determine the sample size. It was revealed that at least four subjects in each group with 8 iterations would provide a power of 0.8 with a two tailed alpha of 0.05. This analysis was based on prior testing with the Immersion CyberGloves® (San Jose, CA). G*Power 3 (http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/, freeware, accessed 2 26 2010) was employed for the calculations.

A cadaveric femur with all soft tissue removed was stabilized in a clamp secured to a table at a height similar to the operating room environment. All participants were given rudimentary instructions to identify what the instruments were and in what sequence to use them. They were shown generally where in the femur to drill the holes. Specific technical instructions were not given. Each participant sequentially drilled a 3.2 mm bicortical hole in cadaveric femoral diaphyseal bone using a Small Battery Drive drill (Synthes®, West Chester, PA), tapped threads into the hole by hand using a 4.5 mm tap on a T-handle (Synthes®, West Chester, PA), and inserted a 4.5 mm cortical screw (Synthes®, West Chester, PA), into the hole until the head was tight against the near cortex. Each participant repeated this sequential task 10 times. Real time wrist, hand and finger position was recorded bilaterally using Immersion CyberGloves® (San Jose, CA) and the Ascension Liberty Tracker® (Milton, VT).

Figure 1 Setup orthopedic drilling.

The CyberGloves® use bend sensing technology to measure angles and magnetic trackers to detect 3-D position and orientation of the hand, fingers and forearm in free space (Figure 1).The kinematics for the wrist and digits were measured. Flexion/extension and abduction/adduction angles were recorded for wrist kinematics. Digit kinematics consisted of: angles at the metacarpal-phalangeal (MCP), proximal and distal interphalangeal (PIP and DIP, respectively) joints of the four fingers and the angle of abduction (ABD) between adjacent fingers. For the thumb, the MCP, ABD, and interphalangeal (IP) angles were measured together with the degree of thumb rotation (ROT) about an axis passing through the trapeziometacarpal joint of the thumb and index MCP joint. The arch of the palm was also measured, with minimum values assigned to the thumb and little finger touching and maximum values correlating with maximal separation [23]. The Ascension Liberty Tracker® captures movements of the hand and the forearm.

Continuous data was collected from the CyberGloves® and Ascension Liberty Tracker® including the angles orientation and position of the fingers, hand, and forearm throughout all 10 repetitions of all 3 tasks for each participant using the software associated with the CyberGloves®. Four skill level metrics were calculated:

1. Task duration (sec). Task Duration was measured from the start of drilling of the first hole to the completion of inserting the 10th screw.

2. Gesture Proficiency. The Gesture Proficiency Score was determined from the algorithm proposed by Kahol et al. [23] based on a hierarchical model of the hand. The human hand can be modeled as a complex system of hierarchically connected rigid units [24]. Each segment can move independently with hierarchical interactions. An algorithm was defined to model hand movements as a function of

activity for individual segments of the hierarchy using segmental Momentum (M), Kinetic Energy (KE) and Force (F), based on velocity and acceleration estimates, and mass calculations. Gaussian low-pass filtering was used to smooth the measure of activity. The dimensionality of the motion sequences was reduced to reveal a manifold space separating experts, intermediate performers and novices using the isomap technique. Complex motion sequences were divided into simpler units employing a gesture segmentation algorithm based on activity profiles [25]. Gesture sequences were compared using dynamic time warping.

Ideal movements for each gesture were determined using OpenSim musculoskeletal 3D models [26] to identify the posture and configuration of the hand that are used to perform certain movements while requiring minimum loading on the hand muscles and skeletal systems. OpenSim musculoskeletal models have been employed by kinesiology and bioengineering community for various purposes like designing ergonomic devices and orthopaedic rehabilitation; however, to our knowledge this is the first study that employs such sophisticated models for determining optimal motion. A CT Scan of a femur was coupled with a virtual drill made in Maya 3D Animation (AutoDesk, San Rafael, CA) and virtual hands from the VirtualSim models (VirtualSim, Coaticook, QC, Canada). The hand movements that are used to produce an ideal hole from any specific starting point were then simulated. These hand movement profiles were treated as optimal motion profiles and used to determine a template of ideal gestures for drilling. Similarly, ideal gesture templates of tapping and screw insertion were determined. Hidden Markov Models were then trained for tool movement analysis employing the Viterbi algorithm [26]. Movements of the participants were then compared to these ideal gestures through the forward backward algorithm [26]. The Gesture Proficiency Score between 0 and 1 is the probabilistic score based on similarity between the ideal movement and movement performed by the participant.

3. Hand Movement Smoothness. Hand movement smoothness refers to the overall smoothness of motion during the task. It is calculated using normalized gross acceleration data of the hand obtained from the CyberGloves® using the formula given below in Equation (1)

$$HandMovementSmoothness = \sqrt{(1 - NormalizedAcceleration)^2} \quad (1)$$

This measure has been validated in previous studies [27].

4. Errors. Errors consist of the drill slipping off of the bone, an inappropriate angle of approach, and readjusting the orientation of the drill beyond the tolerance of a partially drilled hole after starting to drill. All of these errors are important to monitor, because such errors in the operating room are a risk to the patient's safety. Models were developed for slip movements and inappropriate angles of hand and drill. Errors were detected, logged and counted by the motion analysis system.

Learning was determined as improvement in the four performance metrics.

Data was subjected to statistical analysis using ANOVA for skill level metric analysis. The data is parametric and normality was confirmed. The level of significance was defined as $p < .05$.

Study 2

Institutional Review Board approval was obtained through Banner Health and informed consent was obtained from all participants. Four groups of obstetricians/gynecologist participants were recruited based on experience:

1- No Skills (First year residents with minimal to no practice of their suturing skills)

2- Limited Skills (Second year residents with limited practice of their suturing skills);

3- Advanced Skills (Upper level residents with ample suturing practice)

3- Expert Skills (practicing obstetrician/gynecologists with extensive use of suturing as a routine part of their practice).

A total of twenty participants were enrolled; five in the No Skills group, five in the Limited Skills group, five in the Advanced Skills group, and five in the Expert Skills group. Power analysis was performed to determine the sample size. It was revealed that at least four subjects in each group with 8 iterations would provide a power of 0.8 with a two tailed alpha of 0.05. This analysis was based on prior testing with the Immersion CyberGloves® (San Jose, CA). G*Power 3 (http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/ , freeware, accessed 12 31 2010) was employed for the calculations.

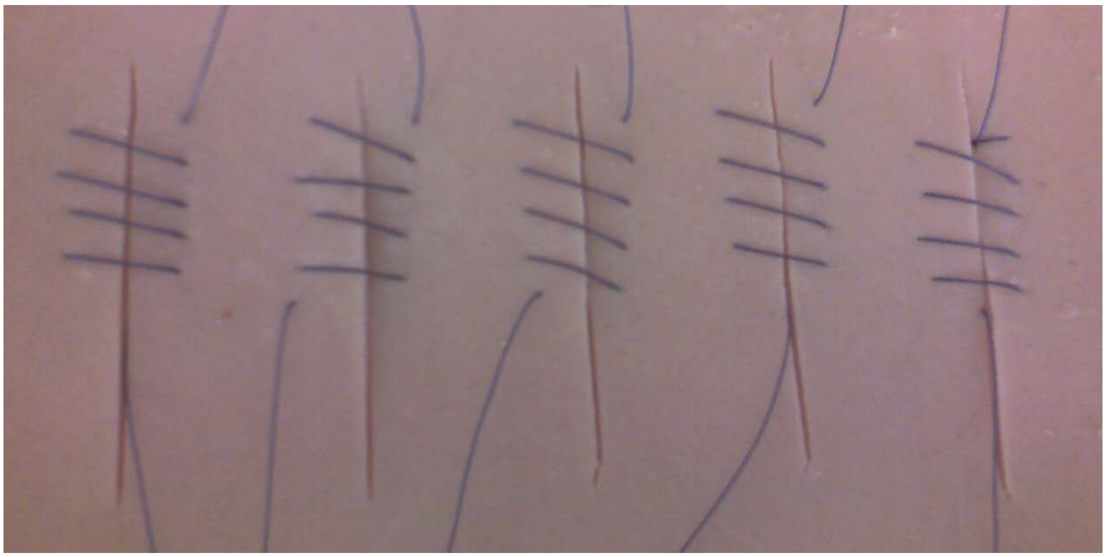

Figure 2 Setup OB/GYN suturing.

An artificial tissue (40 x 20 cm) made primarily of silicone to simulate skin was utilized for placement of suture material. Five incisions of ten centimeters in length were made parallel to one another in the synthetic skin. The artificial tissue was secured on a table at a height similar to the operating room environment. All participants were given rudimentary explanation of the instruments (needle driver, forceps, and suture) and were instructed to place five continuous running sutures across each incision. Specific technical instructions were not given. During each trial, the participants repeated this sequential task 5 times. All participants completed two trials. Real time wrist, hand and finger position was recorded bilaterally using Immersion CyberGloves® and the Ascension Liberty Tracker®. The subjects were videotaped during the sessions in which they completed the tasks and were rated by a skilled physician and the former website www.ratethesurgeon.com based on their basic suturing skills.

Results

Study 1

Data was acquired for four skill level metrics: task duration, gesture proficiency, hand movement smoothness, and number of errors over the course of ten trials for all three participant groups: No Skills, Limited Skills, and Expert Skills. Learning curves (Fig 3,5,7,9) for each participant group and bar graphs for comparison of participant groups (Fig 4,6,8,10) are plotted for each of the four skill level metrics. Learning occurred for all metrics in the No Skills and Limited Skills Groups and for Hand Movement and Task Duration in the Expert Skills Group. Learning in Gesture Proficiency was limited by the ceiling effect for the Expert Skills Group. Statistically significant differences were found for learning in all skill level metrics between all three participant groups ($p<.007$).

Task duration (Fig. 3,4) was similar for both the No Skills and Limited Skills groups, with a mean of 142.6 and 136.1 seconds per trial, respectively. The Expert Skills group had a mean of 107.4 seconds per trial. All three groups demonstrated a decrease in the amount of time to complete the tasks, while the Expert Skills group had the greatest decrease. The total time for completion of all ten trials, including rest time between trials was 1545, 1483, and 1160 seconds for the No Skills, Limited Skills, and Expert Skills groups, respectively. Not only did the Expert Skills group complete the tasks the fastest, but they also required the least amount of time between repetitions.

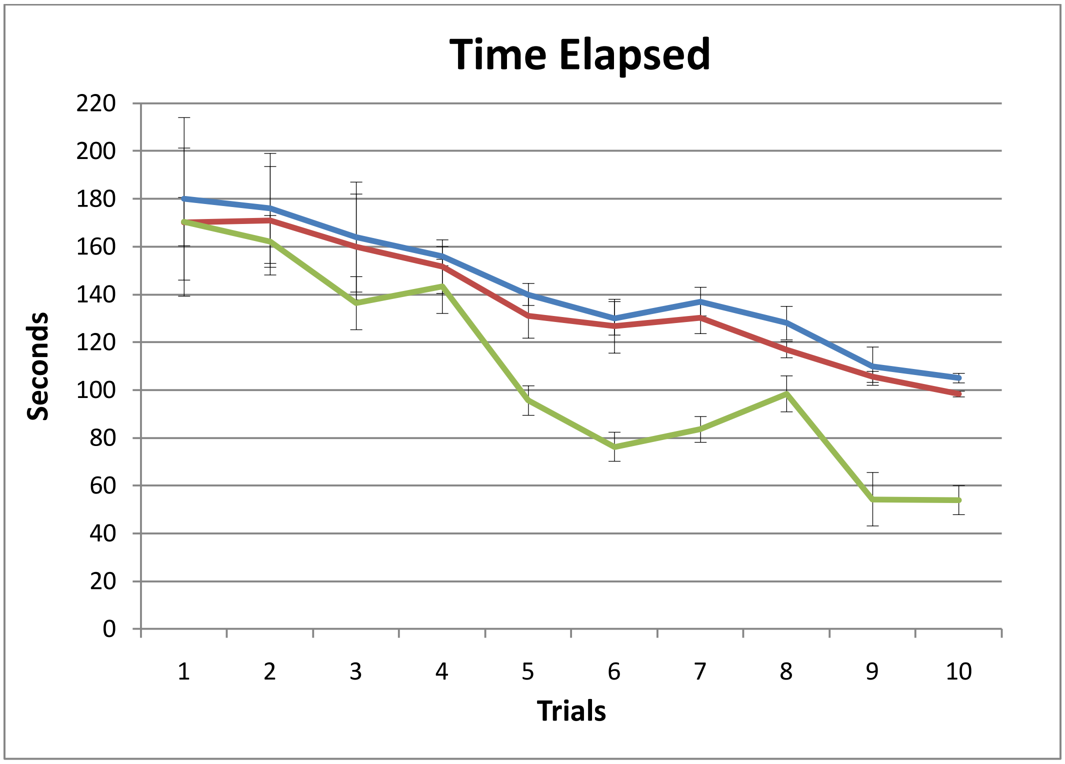

Figure 3. Time elapsed for trials 1 – 10 for each skills group. Blue represents the No Skills group, Red represents the Limited Skills group, and Green represents the Expert Skills group.

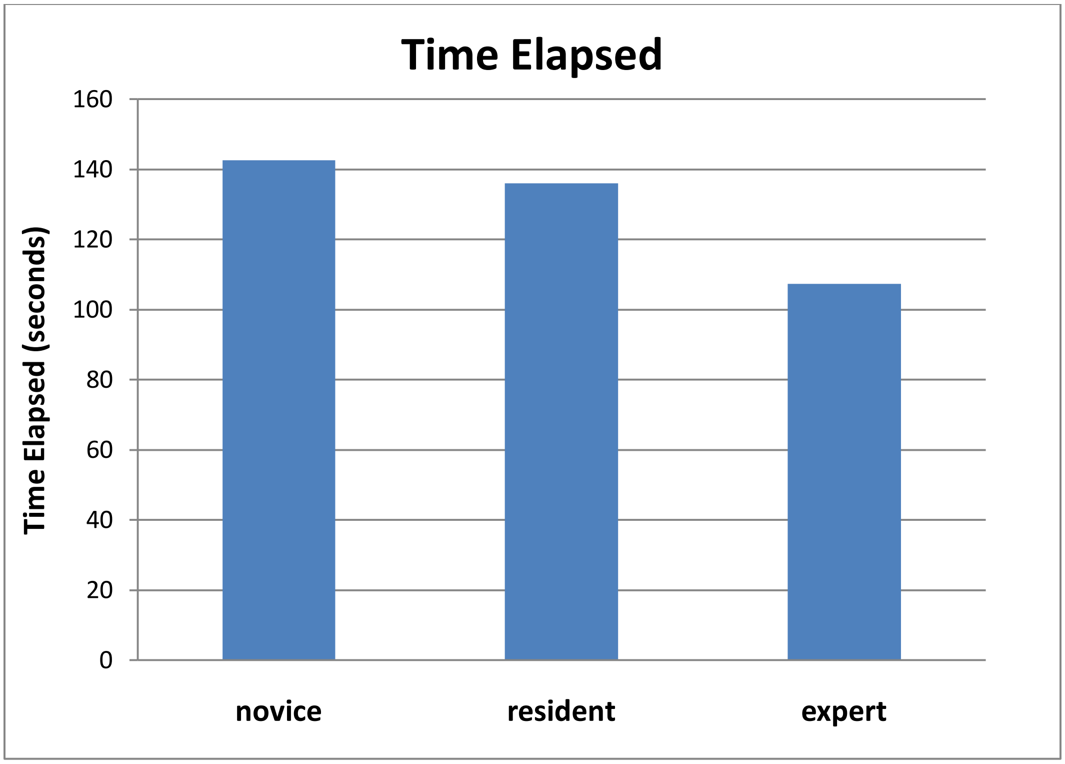

Figure 4 The mean number of seconds over ten trials for each skills group.

Gesture proficiency (Fig 5, 6) demonstrates a similarity between the No Skills and Limited Skills groups. The mean proficiency of No Skills was 0.470 and the Limited Skills was 0.524. The Expert Skills group mean was 0.697. Expert Skills group demonstrated improvement over the first seven trials. Gesture proficiency was approximately 0.90 on the 7th repetition and did not improve further by the 10th repetition, consistent with a ceiling effect. The No Skills and Limited Skills groups made less improvement in gesture proficiency.

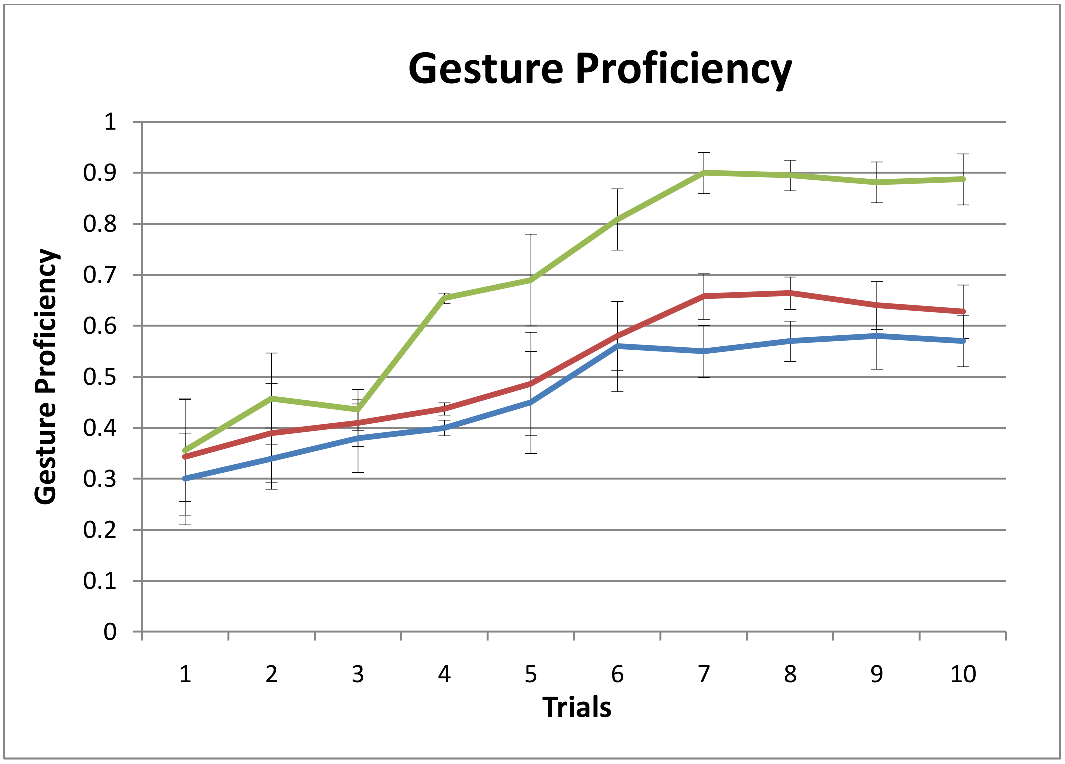

Figure 5 Gesture proficiency for trials 1 – 10 for each skills group. Blue represents the No Skills group, Red represents the Limited Skills group, and Green represents the Expert Skills group.

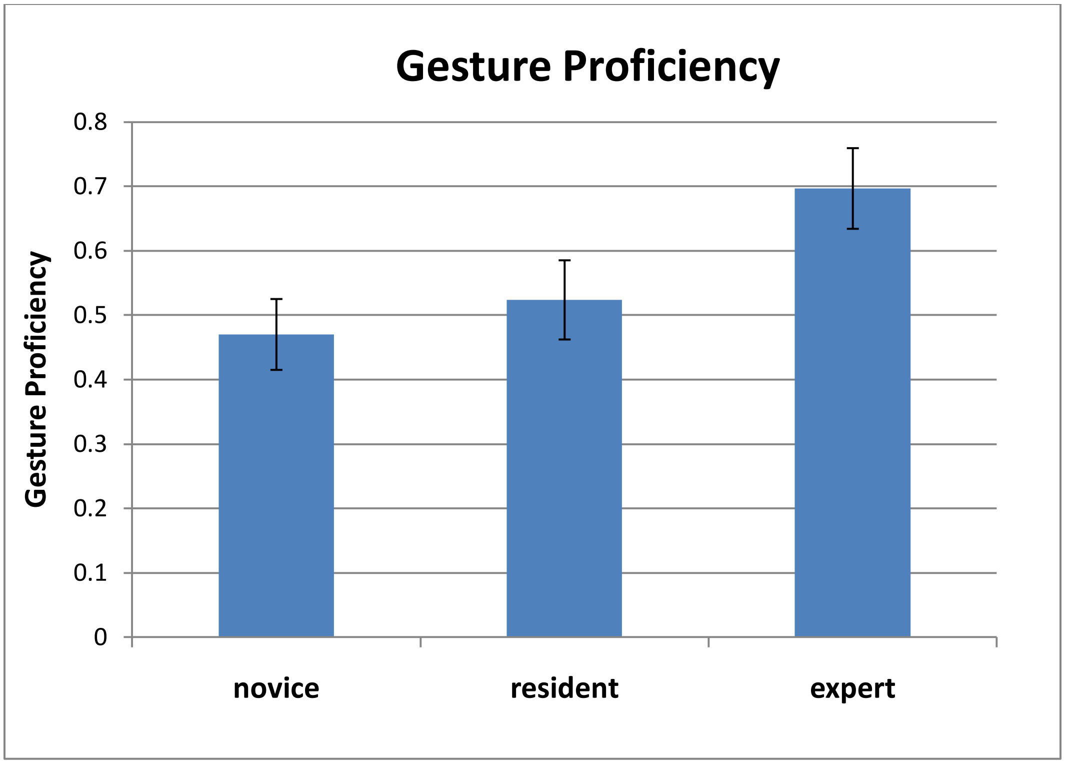

Figure 6 Mean Gesture proficiency over ten trials for each skills group.

Hand Movement Smoothness (Fig 7, 8) data were 0.341 for the No Skills group, 0.383 for the Limited Skills group, and 0.506 for the Expert Skills group. When followed over ten trials, all groups achieved improvement in hand movement smoothness with the Expert Skills group improving the most (0.28 start to 0.82 finish). Hand movement smoothness did not have a ceiling during the 10 repetitions of the tasks studied.

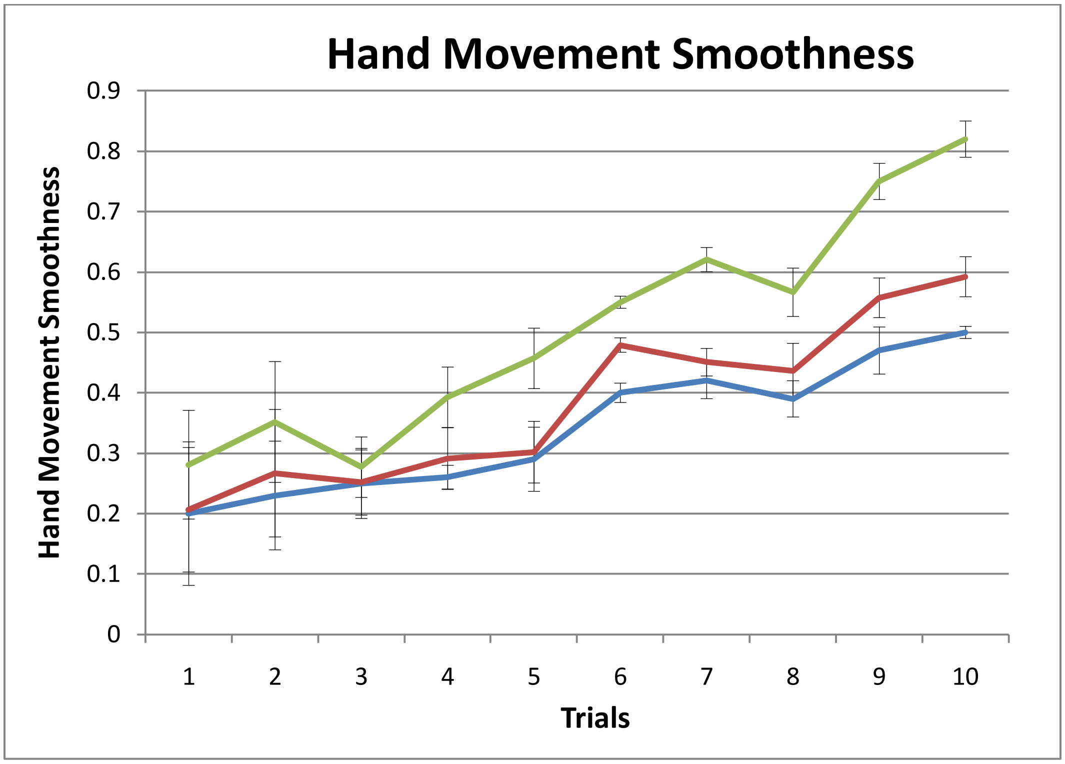

Figure 7 Hand Movement Smoothness for trials 1 – 10 for each skills group. Blue represents the No Skills group, Red represents the Limited Skills group, and Green represents the Expert Skills group.

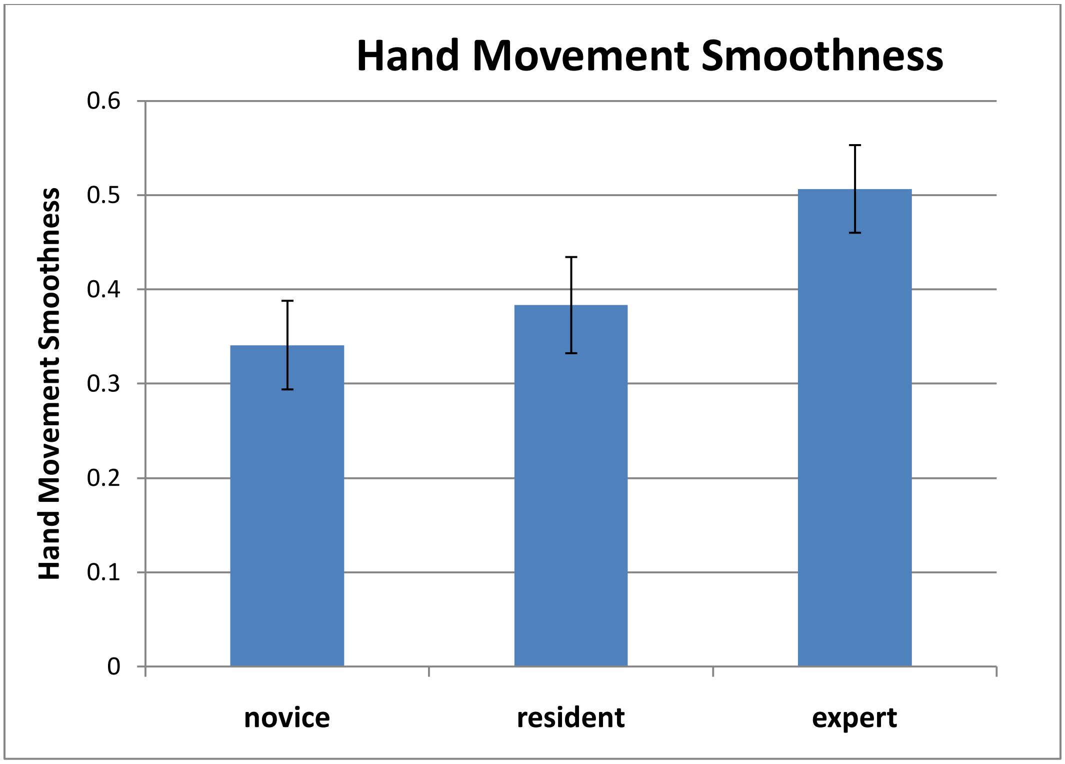

Figure 8 Mean Hand Movement Smoothness over ten trials for each skills group.

Errors (Fig 9, 10) occurred on average 3.2 times per task for the No Skills group, 2.4 for the Limited Skills group and 0.87 for the Expert Skills group. All three groups demonstrated a decreasing rate of errors over the course of ten repetitions. The greatest differentiation between participation groups for errors was found while tapping, as the No Skills and Limited Skills groups had a mean of 1.3 errors per task and the Expert Skills group had no errors. Screw insertion was the least differentiating of the task, as the number of errors for screw insertion was similar in all participant groups.

Figure 9 Number of Errors for trials 1 – 10 for each skills group. Blue represents the No Skills group, Red represents the Limited Skills group, and Green represents the Expert Skills group.

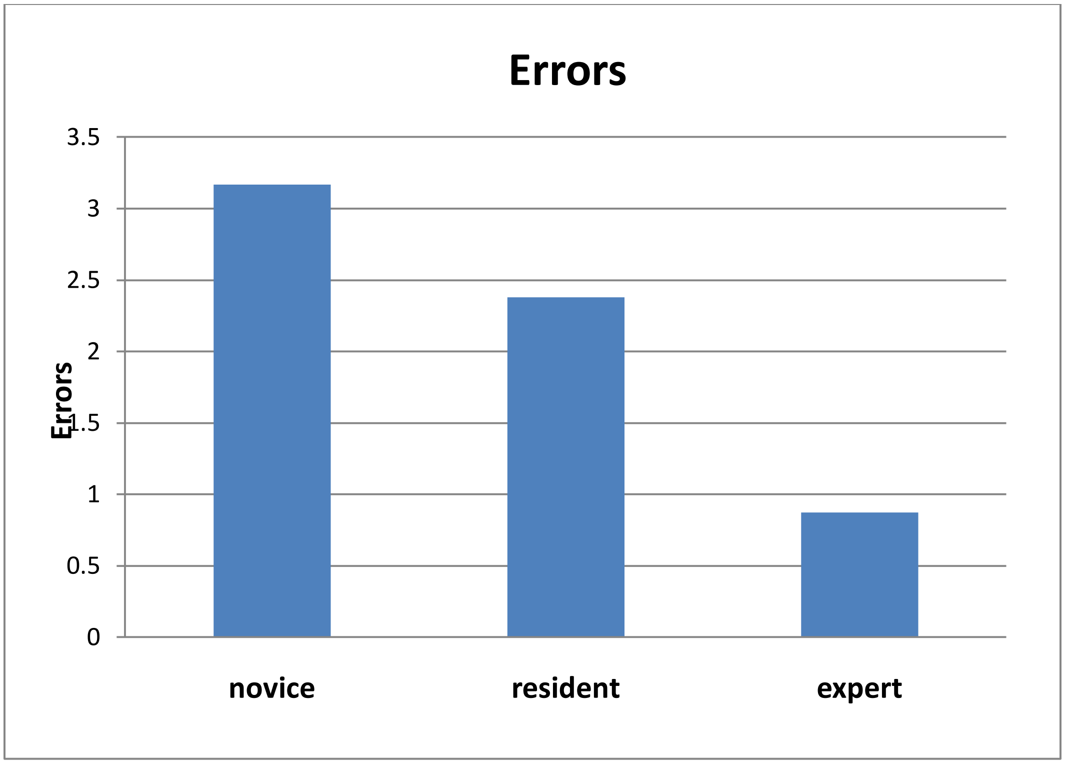

Figure 10 Mean number of Errors over ten trials for each skills group.

Study 2

Data was acquired for four skill level metrics: task duration, gesture proficiency, subjective rating, and number of errors over the course of two trials for all four participant groups: No Skills, Limited Skills, Advanced Skills, and Expert Skills. Bar graphs for comparison of participant groups (Fig 11-14) are plotted for each of the four skill level metrics.

Task duration (Fig. 11) in seconds was recorded for the time elapsed for each trial. The No Skills group required the longest amount of time to complete the tasks, with a mean of 433 seconds on the first trial and 438 seconds on the second trial. The Limited Skills group demonstrated a significant improvement over the No Skills group with a mean of 279 and 246 seconds per trial, respectively. The Advanced Skills and Expert Skills groups both showed significant improvement between trials. The Advanced Skills group performed the trials in 273 and 210 seconds, respectively. The Expert Skills group completed the tasks in trial one in 255 seconds and in trial two in 207 seconds. The Expert Skills group completed the tasks the fastest, as their suturing efficiency is superior to the other three groups.

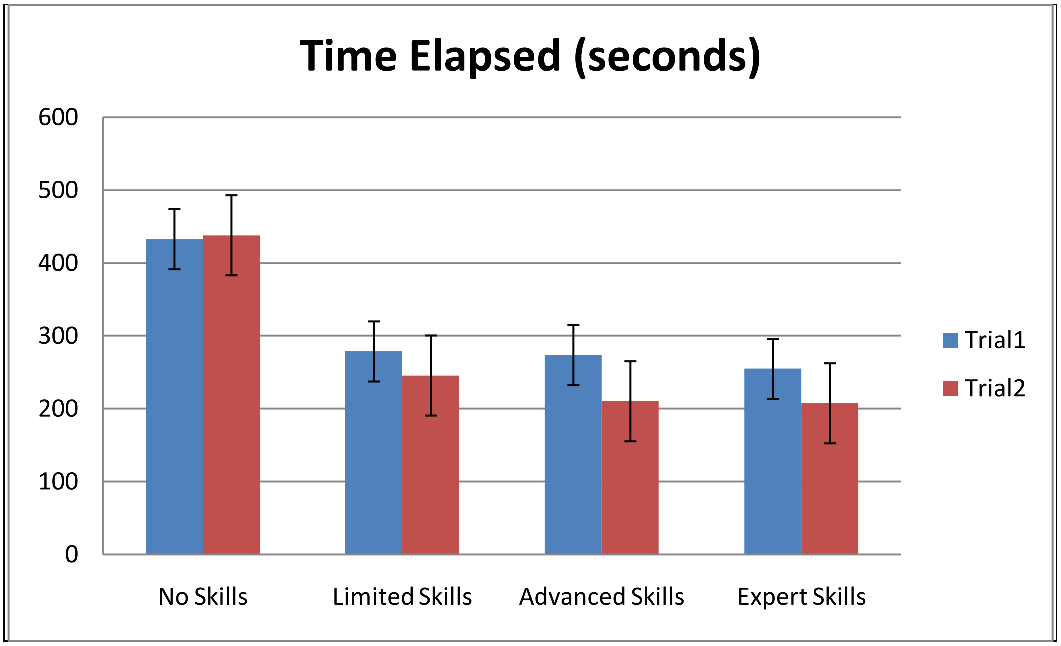

Figure 11 Time elapsed to complete the tasks in each trail for the four skills groups.

Gesture proficiency (Fig 12) demonstrates a difference between all skills groups. The mean proficiency of No Skills was 0.42 and 0.48 respectively. The Limited Skills group's mean proficiency was 0.52 and 0.58 in trials one and two respectively. The Advanced Skills group had a mean proficiency of 0.71 and 0.74. The Expert Skills group mean was 0.74 and 0.81. All groups demonstrated improvement between the first and second trials, with the Advanced Skills group showing the least improvement of the four groups. No group reached maximum proficiency of 1, because it is based off an ideal mathematical model. However, a significant difference in proficiency between the No Skills and Expert Skills groups exists.

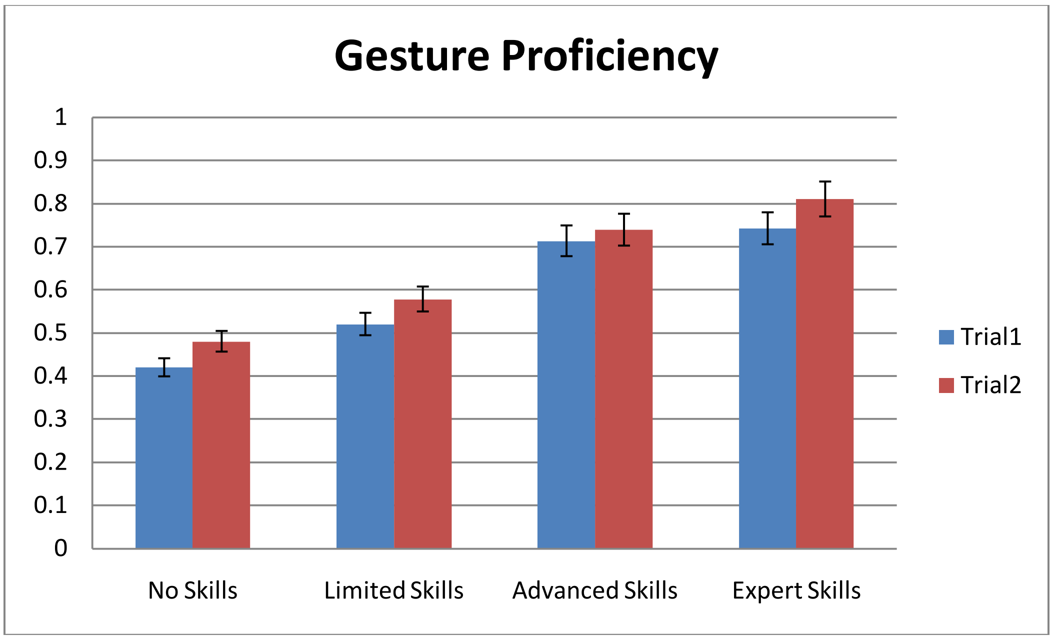

Figure 12 Gesture proficiency achieved by the four skills groups during two trials.

Subjective rating was determined by skilled physicians based upon the videotaped performance of each participant. The video captured each participants' hands as they placed the suture across the five incisions. A rating for each trial is shown in Figure 13, with a scale of 0-10 with 10 being the best. The No Skills group had an average of 3.7 and 4.3 points. The Limited Skills group had an average of 4.6 and 5.6 points. The Advanced Skills group had an average of 7.6

points in two trials. The Expert Skills group had an average of 7.6 and 8.4 points. There is a significant difference between the Expert Skills group and the No Skills and Limited Skills groups, however, the Advanced Skills group nearly achieved the same average rating as the Expert Skills group.

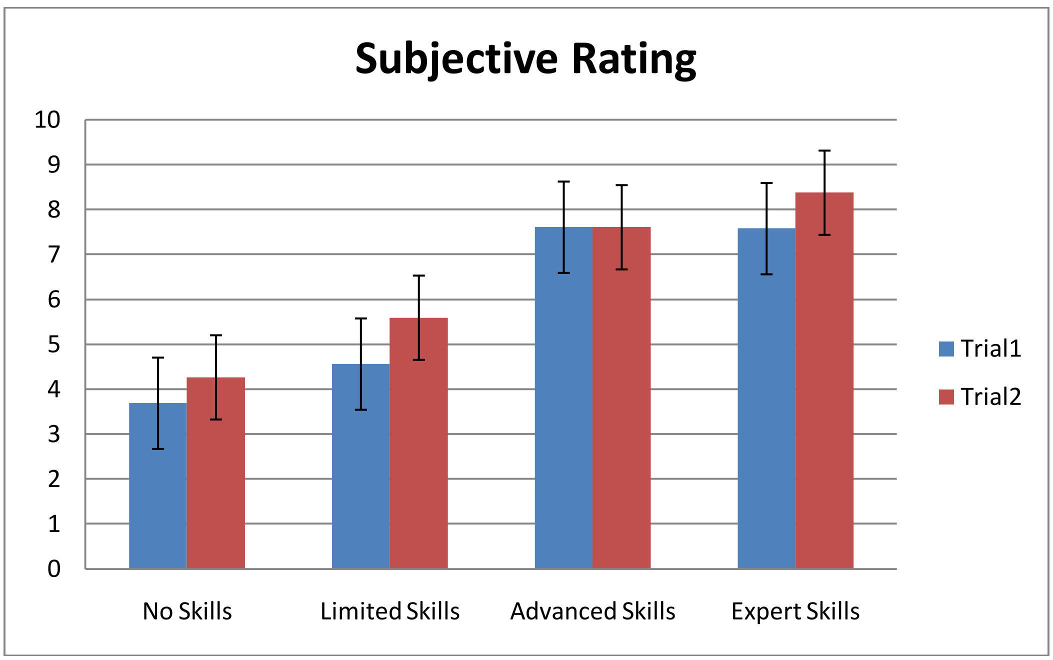

Figure 13 Subjective rating, as determined by skilled physicians based on video footage, of participants' performance in all four skills groups in two trials.

The number of errors were counted for each trial. Errors were defined as fixing the placement of the needle within the needle driver, misplacement of the needle in the tissue, driving the needle more than once from one point of the incision to the other. The No Skills group had an average of 5.8 and 6.5 errors. The Limited Skills group had an average of 4.6 and 5.6 errors. The Advanced Skills group had an average of 2.8 and 1.9 errors. The Expert Skills group had an average of 0.6 and 0.4 errors.

Figure 14 Number of errors in completion of the tasks in two trials for all four skills groups.

The Expert Skills group demonstrated very few errors, a statistically significant difference compared to the other groups. Surprisingly, the No Skills and Limited Skills groups' performances suffered on the second trial, yielding more errors than their first attempt. The Advanced Skills group demonstrated learning as they improved from trial 1 to trial 2.

Discussion

Motion analysis of the hand was performed for four basic surgical skills. Improvement was documented with repetitions of the tasks. Performances of the participants at each level of training were evaluated to determine whether level of training correlated with skill proficiency. Other researchers have performed motion analysis in orthopaedic surgery. Woodrow et al. described a system for measuring wrist movements during spine surgery through electromagnetic trackers and a force plate [18]. That system measures wrist motion, mean forces, peak forces, and task time. For surgeons cannulating a complete set of lumbar pedicles on a synthetic model, a statistically significant difference was seen between experts and novices. A similar study performed by Tashiro et al. analyzed tool motion and force during a simulated arthroscopic procedure [11]. Leong et al. quantified orthopedic skills using experts to subjectively rate performance on video recordings of participants performing tasks on synthetic bone models [19]. Subjective ratings by experts are difficult to assimilate into automated objective quantification systems.

All these studies use tool movements and force profiles as the means of assessing competency of the performer in simulation environments. None of the studies measured hand movement and posture during the task. This is a major limitation of these studies as operator movement and posture are important components of basic motor skills. An algorithm that is based on assessing tool motion and tool forces can objectively analyze proficiency of skills such as drilling; however, it only has limited feedback capacity. Without assessment of hand movements and posture, the determinants of drilling angle and applied force, the feedback will not be as informative to the user. For instance the feedback will describe the incorrect orientation of the tool and which direction/angle it needs to be corrected to for the user to improve the outcome. However, this will not provide the user with precise instructions on how to move their fingers, palm, wrist, or shoulders in order to achieve the desired outcome. Measuring hand movements will allow detailed analysis and feedback, providing corrections for hand, wrist, and shoulder posture and the movements of each to create the desired gestures. Our system is capable of quantifying hand movement and posture in addition to the parameters associated with tool movements which can be estimated by the hand movement profiles.

We capture hand movement data using Immersion CyberGloves®. CyberGloves® have been used to precisely record finger, hand, and forearm position and movement in technical activities ranging from hand gesture computing for the hearing and speech impaired to monitoring and evaluation of virtual bone cement injection [21,22]. Using these gloves, kinematic features of surgeons' hand movements can be analyzed to demonstrate motor proficiency. Previous studies using CyberGloves® gesture recognition have indicated that gesture based analysis of surgical movements may be suitable for analysis of proficiency in general surgery tasks [23]. Gestures are the basic building blocks of tasks. In laparoscopic surgery for example, clockwise rotation and grasping are gesture components of laparoscopic tasks. Complex laparoscopic procedures are segmented into atomic gestures which are then analyzed to reveal proficiency. The overall score in a procedure is the sum of scores for performance of individual gestures and the order in which the gestures are performed. This is a useful method of analysis and is commonly known as task decomposition, as used in OSATS. The advantage is that automatic gesture based analysis of surgical movement including segmenting complex movements into simpler movements can be performed by a computer, eliminating human bias and error.

The participant groups in our study were distinct based on experience in that the No Skills group was neither familiar with the system (Cybergloves, instruments, and individual bone/synthetic skin) nor possessed the task specific motor skills. The Limited Skills group lacked familiarity with the system, yet was familiar with the specific bone site/suturing tools and the task specific motor skills. The Advanced Skills group in Study 2 lacked familiarity with the CyberGloves, yet was familiar with the instruments, suturing material, and possessed the task specific motor skills. The Experts Skills group was fully familiar with the instruments and femoral cortical bone/suturing material, in general, but was not familiar with the CyberGloves® system or the individual bone/synthetic skin they were about to use.

The Expert Skills group had the ability to rapidly adapt the familiar tasks to the experimental bone drilling environment and suturing environment; however, they were more cautious than other groups. Since the experts are already at a very high level of proficiency, one might expect that their group displays the least improvement across the repetitions of the tasks. However, the data demonstrates that the Expert Skills groups improved their speed (as they became less cautious with repetition) and decreased the number of errors made as they adapted to the setup of the system. The setup was different than their typical OR, they were not in the surgical gowns and gloves that they are used to, nor did they have any assistance with tools (such as a scrub tech to hold on to tools not in use) and had to find places for their tools in between uses. Because of these circumstances, it is understandable that they are able to improve more than the other groups as their potential for accuracy and proficiency is much greater than the other groups.

Immersion CyberGloves® can capture objective, quantitative, and continuous performance data from participants performing fundamental orthopaedic tasks as well as gynecologic tasks. The data acquired in these studies was able to objectively distinguish between the skill levels of the participant groups.

Improved performance was detected with increasing task repetitions in all participant groups in Study 1. The Limited Skills group improved the most. This is expected as they begin with a lower baseline score. However, in Study 2, no single group demonstrated superior improvement when assessed across the four metrics. The proportion of system 'familiarization' to motor skills 'learning' was not quantified. Multi-step procedures with decision making were not part of this investigation. Incorporating these considerations into the analysis algorithm will be a component of future work.

When examining the use of time or task duration as an assessment criterion, it is evident that it is weaker in its ability to differentiate between the skill levels. The average time to complete the tasks amongst all groups in the two studies was relatively similar, except for the No Skills group in Study 2. Thus, because this cannot reliably and consistently separate out different skill levels of participants, it has poor construct validity and cannot be used as a stand-alone criterion. It may be possible to develop a system in which the number of errors and the task duration are both utilized in a combined rating system to improve the construct validity.

This gesture proficiency system measures a participant's performance against a theoretical ideal. The main advantage of working with an ideal gesture as the target reference is that the need to identify experts and characterize expert gestures is eliminated. This is very important as often it is hard to determine who actually is an expert and to characterize how they perform gestures. It may be unrealistic to expect that ideal movements can be performed by competent participants; however ideal movements can provide an objective reference to compare performance with. Using this system we can objectively analyze performance of basic skills and identify milestones in skill levels that learners achieve as they become proficient. These milestones will be determined in future research.

Future Directions

While these results are preliminary and need to be validated across larger sample sizes, they do indicate the capability of using CyberGloves® and the associated system for quantitative skills analysis. The main advantage of this system is that one score set is applicable to all types of simulators, tasks, tools and anatomies. Such a system could serve as an important aid in the development of competency driven curriculums and provide benchmarks for basic skills proficiency across tasks and specialties. Procedure proficiency however would require additional procedure specific metrics.

To provide consistency of analysis and to document transferability of skills from training to patient care it is necessary to have similar quantification systems for simulation proficiency and for performance in clinical environments. Our performance metrics are applicable in all simulated environments but the gloves cannot be used during actual patient care. They have been worn with several different simulators and in animal surgery under sterile latex gloves. In order to study hand movement data in the operating environment during transfer studies, refinement of the gloves or alternate technology is needed so that it is unobtrusive and does not violate the sterility of the surgical field. Transfer studies have been performed accepting different performance analysis for simulation than surgery. Howells et al. [20] performed a transfer study where subjects were trained on simulation models and then observed in an OR by expert observers. This showed a statistically significant difference between subjects that did not receive simulation training and those that did. They analyzed movement and forces of tools during simulation and used the Orthopaedic Competence Assessment Project guidelines to assess performance in the OR.

Once further validations are complete, this data set will be used to develop a virtual-haptic simulator to teach orthopaedic motor skills, including suturing. Such a simulator does not currently exist. Further studies will be needed to gather additional data including tissue material properties in addition to other basic orthopaedic skills, such as the use of a saw. In combination with force measurements, from systems such as described by Woodrow et al [18], both kinematic and kinetic information will be used for simulator design. A feedback system based on glove movements will be developed. In the past we have done research that shows a positive impact from providing feedback of real-time hand motion analysis to general surgeons, improving their task proficiency [28]. We will employ a similar system for orthopedic surgery.

Beyond development of a basic skills simulator, we aim to construct a simulator that will challenge the critical thinking skills of its user. Not only will the simulator assess the proficiency of the users' basic skills, but also will assess the cognitive ability of the user while making intra-operative surgical decisions. The simulations will challenge the user to make appropriate decisions for the best outcome of the simulated patient. The literature and expert opinion will be required to come up with the appropriate algorithm for each intra-operative situation and outcome. The resident learners practicing on such simulators will receive valuable feedback that will assist them in recognizing the consequences of their decision making.

In future work, we aim to validate our skill quantification system and lay the foundation for a cross-platform cross-task quantification system for surgical skills. The system will be cross-platform, as the gloves can be worn on any simulator independent of the simulator hardware and in animal labs under sterile gloves. They are cross-task, as we use the task decomposition approach that segments a complex motion into simpler motions that are atomic and form the building blocks of several tasks in orthopaedic surgery.

Conclusions

In conclusion, hand movement data can be acquired for basic surgical skills required to perform tasks such as suturing, drilling a hole, taping threads in a hole and insertion of a screw. That data can be used to identify motor skills levels and rates of learning across skill levels.

References

1. Sourin A, Sourina O, Howe Tet Sen. 2000. Virtual orthopedic surgery training. Computer Graphics and Applications, IEEE; 20, 3, 6-9.

2. Reznick RK. 2005. Surgical Simulation: A Vital Part of Our Future. Anals of Surgery, 242, 5, 640-641.

3. Satava RM. 2004. Disruptive visions: Surgical education. Surgical Endosocopy, 779-781.

4. No authors listed. 2008. The ACGME's Approach to Limit Resident Duty Hours. http://www.acgme.org/acWebsite/dutyHours/dh_achievesum0708.pdf (accessed 27 April09).

5. Fried GM, Feldman LS, Vassiliou MC, et al. 2004. Proving the value of simulation in laparoscopic surgery. Annals of Surg, 240(3):518,25; discussion 525-8.

6. Swanstrom LL, Fried GM, Hoffman KI, Soper NJ. 2006. Beta test results of a new system assessing competence in laparoscopic surgery. J Am Coll Surg, 202(1):62-9.

7. Gardner R, Raemer DB. 2008. Simulation in obstetrics and gynecology. Obstet Gynecol Clin North Am, 35(1):97-127.

8. Naik VN, Matsumoto ED, Houston PL, et al. 2001. Fiberoptic Orotracheal Intubation on Anesthetized Patients: Do Manipulation Skills Learned on a Simple Model Transfer into the Operating Room? Anesthesiology, 95(2): 343-348.

9. Goldmann K, Steinfeldt T. 2006. Acquisition of basic fiberoptic intubation skills with a virtual reality airway simulator. J Clin Anesth 5;18(3):173-8.

10. Soper NJ, Fried GM. 2008. The Fundamentals of laparoscopic surgery: Its time has come. Bulletin of the American College of Surgeons, 93, 9.

11. Tashiro, Y., Miura, H., Nakanishi, Y., Okazaki, K., & Iwamoto, Y. (2009). Evaluation of Skills in Arthroscopic Training Based on Trajectory and Force Data. Clinical Orthopaedics and Related Research®, 467(2), 546-55

12. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. Annals of Surgery 2005; 241: 364-372.

13. Hamstra SJ, Dubrowski A. Effective training and assessment of surgical skills, and the correlates of performance. Surgical Innovation 2005 March 1;12(1):71-7.

14. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. BMJ; BMJ 2003 November 1;327(7422):1032-7.

15. Sidhu RS, Grober ED, Musselman LJ, Reznick RK. Assessing competency in surgery: Where to begin? Surgery 2004 1;135(1):6-20.

16. Dubrowski A, Backstein D. The contributions of kinesiology to surgical education. *J Bone Joint Surg* 2004, 86(12):2778-81.

17. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A. Relationship between skill and outcome in the laboratory-based model. Surgery 2002 3;131(3):318-23.

18. Woodrow, S. I., Dubrowski, A., Khokhotva, M., Backstein, D., Rampersaud, Y. R., & Massicotte, E. M. (2007). Training and Evaluating Spinal Surgeons: The Development of Novel Performance Measures. Spine, 32(25), 2921-2925

19. Leong, J. J. H., Leff, D. R., Das, A., Aggarwal, R., Reilly, P., Atkinson, H. D. E., et al. (2008). Validation of orthopaedic bench models for trauma surgery. J Bone Joint Surg Br, 90-B(7), 958-965.

20. Howells, N., Gill, H., Carr, A., Price, A., & Rees, J. (2008). Transferring simulated arthroscopic skills to the operating theatre: A Randomised Blinded Study. Journal of Bone and Joint Surgery, 90(4), 494-499.

21. Pradhan G, Prabhakaran B, Chuanjun Li. 2008. Hand-gesture computing for the hearing and speech impaired. Multimedia, IEEE, 15(2):20-7.

22. Chui C, Ong JSK, Lian Z, et al. 2006. Haptics in computer-mediated simulation: Training in vertebroplasty surgery. Simulation & Gaming, 37(4):438-51.

23. Kahol K, Krishnan NC, Balasubramanian VN, et al. 2006. Measuring movement expertise in surgical tasks. In Proceedings of the 14th annual ACM international conference on Multimedia:. Santa Barbara 23-27 October 2006. 2006: 719 – 722.

24. Jones, L. A., & Lederman, S. J. (2006). Human Hand Function. New York: Oxford.

25. Kahol, K., Tripathi, P., & Panchanathan, S. (2003). Gesture Segmentation in Complex Motion Sequences. Proceedings of the IEEE International Conference on Image Processing, Barcelona, Spain. II, 105-8.

26. Delp, S. L., & Loan, J. P. (1995). A software system to develop and analyze models of musculoskeletal structures. Computers in Biology and Medicine, 25(1), 21-34 .

27. Kahol, K., Smith, M., Mayes, S., Deka, M., Deka, V., Ferrara, J., et al. (2007). The Effect of Fatigue on Cognitive and Psychomotor Skills of Surgical Residents In Foundations of Augmented Cognition (Vol. 4565, pp. 304-313). Berlin / Heidelberg: Springer.

28. Kahol, K., Smith, M., & Tripathi, P. (2006). Gesture Based Hand Movement Analysis and Haptic Feedback for Surgical Training. Proceedings of Medicine Meets Virtual Reality 14, Long Beach, CA.

29. Goff, B. A., Lentz, G. M., Lee, D., Fenner, D., Morris, J., & Mandel, L. S. (2001). Development of a bench station objective structured assessment of technical skills. *Obstetrics & Gynecology, 98*(3), 412-416.

30. Lentz, G. M., Mandel, L. S., Lee, D., Gardella, C., Melville, J., & Goff, B. A. (2001). Testing surgical skills of obstetric and gynecologic residents in a bench laboratory setting: Validity and reliability. *American Journal of Obstetrics and Gynecology, 184*(7), 1462-1470.

31. VanBlaricom, A. L., Goff, B. A., Chinn, M., Icasiano, M. M., Nielsen, P., & Mandel, L. (2005). A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS). *American Journal of Obstetrics and Gynecology, 193*(5), 1856-1865.

Objective Methodology to Quantify Motor Skills in Basic Orthopaedic and Gynecologic Surgical Tasks

THE UNIVERSITY OF ARIZONA COLLEGE OF MEDICINE PHOENIX

Sarah Kennedy Sharer; Kanav Kahol Ph.D, Arizona State University;
Matthew Nugent M.D., Alex McLaren M.D., Melissa Gioia M.D. Banner Good Samaritan Medical Center

Introduction

- Skill level of surgeons is an important metric to capture in a training environment.
  - To date, teaching of skills has been subjective.
- Objective outcome measures in surgery are identified in the literature, and are effective in differentiating levels of expertise in surgical tasks.
  - An experienced surgeon demonstrates expertise through sophisticated motor skills, which are refined, smooth, and very efficient.
  - Comparatively, a novice's performance appears awkward and highly inefficient.
- Fine level detailed analysis in surgical procedures reveals important mechanisms involved in surgical learning.
- There is a need to develop a cross-compatible evaluation system that can be employed with many simulators, skills, and even in the OR.

Hypothesis

Surgical skill levels of individuals with various years of training can be differentiated using data gloves (Immersion CyberGloves ®).

Materials & Methods

Study 1 – Orthopaedic Drilling

IRB approval through Banner Health and informed consent from all participants. 16 Participants were recruited based on experience:

- No Skills - 6 medical students that had never used a drill
- Limited Skills – 4 junior orthopaedic residents with 6 months – 1 year of exposure
- Expert Skills – 6 practicing orthopaedic surgeons with extensive use of these skills

Figure 1: Setup orthopaedic drilling

Task:
Drill a 3.2 mm hole in cadaveric femur, tap threads into the hole using a tap on a T-handle, and insert a 4.5 mm cortical screw into the hole until the head was tight against the near cortex. Repeat 10 times. Real time wrist, hand and finger position was recorded using Immersion CyberGloves® and the Ascension Liberty Tracker®.

Study 2 – Gynecologic Suturing

IRB approval through Banner Health and informed consent from all participants. 20 participants enrolled:

- No Skills – 5 First year residents with minimal to no practice of their suturing skills
- Limited Skills – 5 Second year residents with limited practice of their suturing skills
- Advanced Skills – 5 Upper level residents with ample suturing practice
- Expert Skills – 5 Practicing obstetrician gynecologists with extensive use of suturing as a routine part of their practice

Figure 2: Layout Study 2 suturing.

Task:
Place five continuous running sutures across each of the five 10 cm incisions in an artificial silicone skin. Repeat five times. Real time wrist, hand and finger position was recorded bilaterally using Immersion CyberGloves® and the Ascension Liberty Tracker®.

Discussion

Published studies discuss the use of tool movements and force profiles as the means of assessing competency of the performer in simulation environments. None of the studies measured hand movement and posture during the task, important components of basic motor skills. Measuring hand movements allows detailed analysis and feedback. Our system is capable of quantifying hand movement and posture in addition to the parameters associated with tool movements.

Immersion CyberGloves® can capture objective, quantitative, continuous, performance data from participants performing fundamental orthopaedic tasks such as drilling a hole, tapping threads in a hole, and inserting a screw into the tapped hole; as well as gynecologic tasks like placing the needle in the needle driver, driving the needle into the appropriate layer of tissue repeatedly to bring the edges of the wound together. The data acquired in these studies was able to objectively distinguish between the skill levels of the participant groups.

Results – Study 1

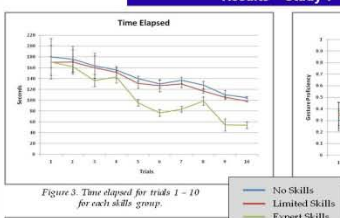

Figure 3. Time elapsed for trials 1 – 10 for each skills group.

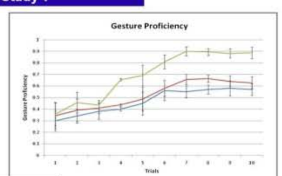

Figure 4. Gesture Proficiency for trials 1 – 10 for each skills group.

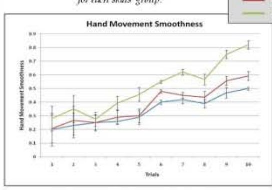

Figure 5. Hand Movement Smoothness for trials 1 – 10 for each skills group.

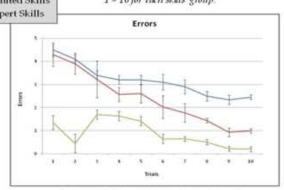

Figure 6. Number of Errors for trials 1 – 10 for each skills group.

Results – Study 2

Figure 7. Time elapsed to complete the tasks in each trail for the four skills groups.

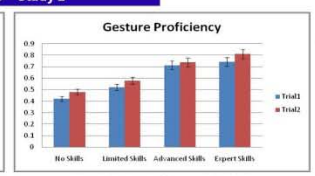

Figure 8. Gesture Proficiency achieved by the four skills groups in two trials.

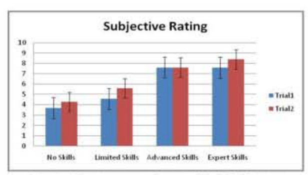

Figure 9. Subjective rating, as determined by skilled physicians based on video footage, of participants' performance in all four skills groups in two trials.

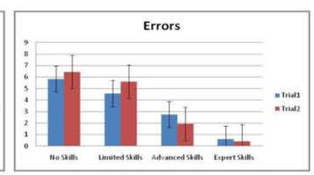

Figure 10. Number of Errors in completion of the tasks two trails for all four skills groups.

Conclusion

In conclusion, hand movement data can be acquired for basic surgical skills required to perform tasks such as suturing, drilling a hole, taping threads in a hole and insertion of a screw. That data can be used to identify motor skills levels and learning across skill levels.

Acknowledgements

The author would like to thank Human Machine Symbiosis Laboratory at Arizona State University and the SimET Center at Banner Good Samaritan Medical Center