The Effect of Blocked Versus Random Task Practice Schedules on the Acquisition, Retention, and Transfer of Surgical Skills

By

Justin Rivard

A Thesis Submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Surgery

University of Manitoba

Winnipeg

Copyright © 2012 by Justin Rivard

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank Dr. Jason Park for taking on the role of supervisor and mentor throughout this project. Without his guidance, support, and insight this project would not have been possible.

I would also like to thank Dr. Amir Hossein and Dr. Bertram Unger for their invaluable assistance and guidance in the design and implementation of the MARS hand-motion analysis system.

I extend my appreciation to all the medical students and residents that took time out of their schedule to participate in this study.

Finally, I would like to thank Christine, Abigail, my mother, and my father for their continued support and patience throughout this entire project.

Abstract

Background: How to optimally integrate simulation into a surgical training program is relatively unknown. We studied the effect of varying the practice schedule into either blocked or random patterns (termed contextual interference) on the long-term retention and transfer of surgical skills.

Methods: 36 participants were randomized to practice 4 tasks from the Fundamental of Laparoscopic Surgery (FLS) program using one of three training schedules (blocked, random, no training). Skill was assessed using FLS scoring and hand-motion efficiency scores.

Results: A positive benefit of training was seen over the controls for all 4 tasks ($p<0.05$). No difference was seen between the blocked and random groups in the amount of skill acquired, skill learned, or transfer of skill.

Conclusion: The application of contextual interference was unable to differentiate between the blocked and random training groups. This could be due to the complexity of the tasks and/or the inexperience of the learners.

Chapter 1. Introduction

Surgical technical skills have traditionally been taught in the operating room. However, a great deal has changed since Halsted's description of a surgical residency in 1904[1]. The Accreditation Council for Graduate Medical Education (ACGME), which is the professional body responsible for accrediting residency programs in the USA, implemented an 80-hour resident workweek restriction in 2003[2]. There is increased emphasis on operating room efficiency, which contributes to decreased time spent towards intraoperative teaching. In addition, growing concern exists from the public on the idea of residents "learning" on real patients[3]. All these things in combination result in decreased opportunity for trainees to gain the experience necessary to maximize their true potential. As a result there has been increased interest in the use of formal basic skills laboratories involving models and simulation during surgical residency, with an overall aim to increase surgical skill and experience in the operating room[4, 5].

The ACGME Residency Review Committee for Surgery (RRC-S) now mandates that simulation be part of residency training[6] and the American College of Surgeons (ACS) has recently launched a program to accredit surgical skills laboratories in the United States and Canada[7]. However, the best way to implement simulators into surgical training remains relatively unknown. In order to teach and assess basic laparoscopic skills, the ACS and the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) have endorsed the bench model simulator-based Fundamentals of Laparoscopic Surgery (FLS) program[8]. Many institutions in North America now require that practicing surgeons complete FLS training prior to performing laparoscopic procedures on patients, and most recently, the American Board of Surgery (ABS) determined that passing the FLS certifying examination will be a requirement for taking the ABS qualifying (written) examination[8].

Fundamentals of Laparoscopic Surgery (FLS) is a two-part program consisting of an online didactic portion and a technical skills portion designed to teach and assess the fundamental knowledge, judgment and technical skills required in basic laparoscopy. The didactic component is based on several components including preoperative considerations, intraoperative considerations, basic laparoscopic procedures, and postoperative care and complications[9]. The technical skills portion incorporates a portable trainer box to facilitate the manual skills component of FLS. The manual skills training practicum was based on the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills program (MISTELS), which was developed and validated by Dr. Fried and colleagues at McGill University in Montreal, Quebec[10-12].

Despite the popularity and widespread implementation of the FLS program, there has been little systemic study on how the FLS program and its individual components are best applied within the context of broader technical skills curricula. The technical skills portion of FLS consists of five tasks with the training protocol instructing learners to repeatedly practice a single task until they achieve a predefined criterion level[13, 14]. Once this level is achieved, the learner proceeds to the next task. It is unclear whether this FLS training protocol is optimal for learning laparoscopic skills.

A program or course that consists of multiple discrete tasks can be practiced in a blocked or random fashion. A blocked practice schedule is one in which a single task is practiced repeatedly before moving on to the next task. In contrast, random practice schedules involve performing the same number of repetitions of each task, but in a random order such that a given task is never practiced on successive trials. In the motor learning literature, Shea and Morgan previously assessed blocked and random practice schedules in a study that required participants to knockdown barriers in a predefined sequence, with each sequence representing a different task[15]. On the immediate post-tests there was a clear advantage found for the blocked practice group. However, on the delayed retention tests given ten days later, there was an advantage for the random practice group regardless of whether the retention tests were given in a random or blocked order.

The findings of Shea and Morgan refute the assumption that the condition that speeds skill acquisition is also most effective for skill retention. Repeated practice in a blocked fashion may help participants to acquire motor skills more quickly, but they do not necessarily retain these skills over time. Conversely, participants who follow random practice schedules may acquire skills more slowly, but may be better equipped to retain their skills over time. The authors hypothesized that the random group was forced to use multiple processing strategies during the acquisition phase, whereas the blocked group were not required to use such multiple processing strategies[15]. This can be viewed as being forced to forget and then subsequently retrieve the information for the tasks multiple times during practice. It could be the exercise of retrieving information for the random group that slows performance during acquisition, but actually confers an advantage during retention and transfer tasks.

While studies in motor learning literature demonstrate an advantage for long-term learning under random practice conditions, the applicability of these findings to surgical training is not entirely clear. Surgery is more complex and requires more cognitive involvement than other motor skills and the principles that apply to learning surgical skills therefore require further study[16, 17]. Dubrowski *et al*. previously deconstructed an orthopedic bone-plating procedure into a series of five smaller tasks and then had surgical novices practice these tasks in either blocked or random order[17]. On retention tests, they did not find any significant differences between blocked and random practice schedules on learning. It should be noted, however, that the retention tests took place only thirty minutes after practice. It is generally believed that the actual duration of the retention period is skill and context dependent, but that it can take many days for the temporary influences of practice to dissipate in order to get a true measure of learning[18, 19]. Brydges et al. used a similar study design but used delayed transfer tests on cadaveric models to assess learning[20]. Their findings again suggested no significant differences between blocked or random practice schedules on skill transfer. However, the FLS training program differs from bone-plating because it consists of multiple independent tasks for which there is not a functional order, which perhaps makes it more suitable for blocked versus random practice schedule comparisons. Furthermore, the individual FLS tasks are arguably more complex than the deconstructed bone-plating skills.

The following chapters will explore the background and knowledge on the current state of surgical skills curricula. As surgical skills courses are synonymous with simulators, it would be impossible to have a meaningful discussion about surgical skills training without introducing what is currently known about simulators. After describing the role of simulation, the discussion will focus on the foundation behind the current recommendations for practice strategy. Finally, since the ultimate goal of any skills curriculum is long-term learning, it is necessary to introduce the theory and various methods that have been studied to allow the retention of motor skills.

Chapter 2. Proven Utility of Simulators

Developing technical skills is essential to surgical residency training. However, the traditional method of learning in the operating room may be inefficient, costly, and may increase patient morbidity[21, 22]. An increasingly popular alternative is to use lab-based simulation, which for laparoscopic skills training can include lower fidelity video-trainers and/or higher fidelity virtual reality trainers. One of the main goals of simulation is for the trainee to practice skills in a safe environment before refining them in the real world. However, the development of effective skill-evaluation metrics is an ongoing challenge.

A growing body of literature has shown that practice in lab-based settings does indeed lead to measurable improvements in surgical skill in the simulated environment and in the operating room, all of which supports the use of surgical simulation[12, 23-29]. Initial research sought to determine whether practice with a simulator would improve surgical skills in a simulated environment. Derossis *et al*. used the original MISTELS tasks and had surgical residents randomized to a weekly practice group or no practice[29]. Their results showed that with weekly practice sessions a significant correlation existed between the trial number and the performance score in each task. Additionally, after 7 trials no plateau was found, indicating that with ongoing practice there is the potential for ongoing learning.

Fairly early on as simulators were in the process of being popularized, it was difficult to evaluate the transferability of skill learned in the simulation lab to outside the simulation lab. With the development of validated tools to evaluate technical skill[30-32], acquiring evidence on the transferability of skill became possible. Perhaps the most studied and commonly used tool to measure technical skill is the Objective Structured Assessment of Technical Skill (OSATS), developed at the University of Toronto by Reznick and colleagues. (OSATS) is a performance-based assessment of technical skills in which expert observers grade the performance of trainees using both a global assessment and checklist[30, 32]. The format of both assessment forms are such that the observers require very little training in order to correctly use them. Observers grade the performers on both a 'yes/no' checklist specific to the task and a global assessment form with seven dimensions of operative performance. The seven dimensions used are 1) respect for tissue 2) time and motion 3) instrument handling 4) knowledge of instruments 5) use of assistants 6) flow of operation and forward planning 7) knowledge of specific procedure. The above dimensions are rated using a likert scale of 1 through 5, with scores of 1, 3, and 5 anchored with descriptors.

Structured grading techniques attempt to standardize evaluation through rated checklists and score sheets. The OSATS[30, 32], was based on the successful format of the Objective Structured Clinical Examinations (OSCE)[33]. The OSATS and OSCE share the same strengths and weaknesses. The grading is typically done in real time via task-specific checklists and/or global assessment charts. These stations and checklists increase standardization of clinical skills assessment. However, they are labour intensive and require prolonged time commitments from faculty to act as expert raters.

Automated and objective assessment methods would eliminate the subjective nature of structured human assessment and free up faculty resources. Computational analysis of surgical motion acquired using a data collection system has shown great promise as a basis for objectively and cost-effectively measuring skill. Movement tracking can be based on electromagnetic, mechanical, or optical systems. The most commonly used system in the surgical literature is the Imperial College Surgical Assessment Device (ICSAD) developed by Darzi and colleagues at the Imperial College of London[34, 35]. The ICSAD system has been shown to have construct validity using open surgical simulation[36] and has also been validated in laparoscopic simulation[34]. Previous studies have shown both ICSAD and OSATS to be effective objective methods of appraising technical skill. Darzi's group have shown that there is a significant relationship between motion analysis assessments, as measured by ICSAD and the OSATS global ratings score[37]. A significant correlation has also been shown between FLS scores and motion efficiency metrics of total distance, number of movements, and total time[38]. Thereby also showing concurrent validity of the ICSAD system.

In one of the first attempts at measuring the generalizability from the bench model to the human model Anastakis *et al*. randomly assigned first year residents into either training on a cadaver model, training on a bench model, or independent reading from a text book[39]. Using the validated OSATS checklist and global ratings scores they found that both the cadaver trained and bench-trained groups outperformed the text-learning group when performing a set of surgical tasks on cadavers. Additionally, there was no difference in the bench and cadaver groups.

Following the success of proving that simulator training improves skill in the human cadaveric model, more researchers started testing this hypothesis on live human patients. Using the Southwestern minimally invasive video trainer, Scott *et al.* randomized junior residents into ten 30-minute training sessions or no training[26]. Using the OSATS global rating scale to score the residents on their performance of a laparoscopic cholecystectomy both before and after training, the authors were able to show a statistically significant difference in improvement between the trained and untrained residents in 4 of the 8 OSATS categories.

Using a virtual reality simulator (MIST VR) Seymour *et al*. studied the effects of training on a single diathermy task until an expert-level proficiency was achieved[27]. Surgical residents were randomized into either training on the MIST VR simulator or no training followed by completion of a laparoscopic cholecystectomy on a live patient in a clinical setting. The conclusion of the study was that residents that trained on the VR simulator performed the laparoscopic cholecystectomy faster, and with fewer errors than the control group. This study was the first to validate the role of VR training on the ability of surgical residents to perform an operative procedure with an improved performance.

With both video trainer and virtual reality simulators showing improvement in surgical skill, speculation began as to whether one method of simulation was superior to the other. In a recent Cochrane systematic review[40], it was reported that the effectiveness of VR simulation for laparoscopic surgery was at least as effective as VT simulation, but that no true clear advantages were seen. Despite the widespread adoption of VR simulators into many surgical residency programs, not all studies show conclusive evidence for transfer of skill or predictive validity[41, 42]. Predictive validity refers to the extent to which test results in one setting correlate or predict performance in another setting. Ideally simulators that have been shown to have predictive validity would be incorporated into training programs. In addition, multiple types of VR simulators exist, each with different levels of realism, degrees of haptic feedback, and cost, all of which makes the decision as to which VR simulators to incorporate into a residency training program all the more difficult.

Chapter 3. Existing Practice Strategy Knowledge

Surgical simulation is rapidly being incorporated into residency programs. However, little is known about how best to implement them into a broader training curriculum. The association of Program Directors in Surgery (APDS) together with the ACS comprises the Surgical Skills Curriculum Task Force. The objectives of this task force are to improve resident performance through skills practice and using assessments of skill to determine "operating room readiness"[6]. The curriculum, which includes skills such as suturing, airway management, stapled and hand-sewn anastomosis, and both basic and advanced laparoscopic skills, proposes that learners will review a video of expert performance, go through guided practice until performance meets standards, and then make a video of their performance to be graded by experts to determine operating room readiness. The method of training used to achieve the predetermined standards is not stated.

One of the first attempts to answer the question "what curriculum is best" for implementing simulators into surgical training was a study by Scott *et al.* They looked at medical students and second and third year surgery residents before and after training on five video trainer tasks[43]. All participants trained for ten 30-minute sessions. The authors concluded that training should include 30-35 repetitions of each task. This recommendation is based upon the fact that the amount of repetitions required to achieve a 90th percentile score was 32. However, the ultimate score was defined as the best score for each group within the 10 training sessions. When looking at the plotted learning curves it is obvious that had the groups been allowed to practice for longer, the ultimate scores (time taken to complete the task) would have been much lower, thereby increasing the recommended minimum number of training repetitions.

Setting the level of proficiency as a predetermined number of repetitions has been shown to be inadequate. Brunner *et al.* looked at second year medical students performing 30 repetitions of various tasks on a computer-based simulator[44]. They found that multiple plateaus occurred for every task studied and that the best score was detected beyond the highest plateau for 9 out of 12 tasks, suggesting that with continuing practice participants may have reached even higher skill plateaus beyond 30 repetitions. In addition, there were large standard deviations in scores even at 30 repetitions, indicating significant variability in individual ability. This implies that a curriculum based upon an arbitrary duration or number of repetitions may be inefficient since some trainees will be over-trained and kept for longer than necessary. Similarly, some participants will be undertrained unless additional practice is afforded.

Feldman *et al.* assessed learning curves for a simple laparoscopy task[45]. In their study they had medical students perform 40 repetitions of the FLS pegboard transfer task. They defined the learning plateau as the theoretical best score achievable, which was calculated by fitting the data to an inverse curve. Overall, subjects were able to reach 90% of the learning plateau in 6 trials. They further separated the data into participants with an interest in surgery and those without. Those interested in surgery reached 90% of the learning plateau in 4.6 trials whereas it took 8.2 trials for those not interested in surgery. However, the sample sizes were small prohibiting any firm conclusions. Differences in learning plateaus were reflected by the degree to which the scores were continuing to improve after 30 or 40 attempts. This again shows that some subjects may require longer practice to reach their full potential.

The literature is prevalent with studies showing that acquisition curves are different for different learners. With some individuals attaining skill and learning faster than others, the question remains as to what is the best way to objectively assess when a learner has completed their training. One type of performance standards is to measure expert levels on a particular type of simulator. Then trainees can practice until they meet these same expert levels, thereby rendering them as "proficient" as experts in that particular task or simulator. This is where the term proficiency training comes from. Several studies have looked at training individuals to proficiency standards and then comparing those groups to individuals not training to proficiency. One such study was performed by Seymour *et al.* where they had surgical residents in PGY 1-4 train on a single diathermy task as part of a virtual reality simulator[27]. Both the proficiency group and the control group had similar baseline characteristics. Following training or no training, a blinded intraoperative dissection of the gallbladder off of the liver bed as part of a laparoscopic cholecystectomy was assessed. The proficiency-trained group performed the procedure faster and with six times fewer errors than the control group.

As part of the ongoing research efforts for the FLS program, Scott *et al.* trained 21 medical students to expert-derived performance goals using a previously developed proficiency-based FLS skills curriculum[14, 46]. Each task was practiced in order (tasks 1-5) until proficiency was reached, or a maximum number of 80 repetitions were reached. Proficiency was achieved for 96% of the five tasks during training and required an average training time of approximately 10 hours. Their results showed that at baseline no trainees achieved a passing score, while, a 100% certification pass rate was achieved following proficiency training. This study by Scott *et al*. is an example of a blocked training schedule. While this shows the success of blocked training to proficiency, it remains unclear whether this is the best method for long-term learning. An additional strength of this study is that by showing success of the curriculum in medical students, the curriculum will likely be suitable for residents or practicing surgeons regardless of their prior levels of experience.

Arguably, the main benefit to proficiency training is that all trainees should uniformly attain the same benchmark level of performance[47]. Additionally, subjects can learn at their own pace, minimizing unnecessary repetition in those who acquire the skill more quickly, while ensuring adequate training in those with slower skill development or less innate ability. Proficiency training allows for deliberate practice. Deliberate practice refers to a form of training that involves constant, repetitive, goal-oriented practice where the learner continuously monitors their performance and subsequently corrects, experiments, and reacts to immediate and constant feedback, with the aim of constant improvement[48]. Deliberate practice consistently leads to improved performance among trainees[49] and is likely a major contributor to the success of the proficiency training model.

Another factor that can affect the rate of skill acquisition and the ultimate level of learning is the distribution of practice[47, 50]. The spacing of practice sessions can either occur as one long session (termed mass practice) or multiple short practice sessions (termed interval practice). Although both study groups show measurable performance improvements, the interval practice group routinely shows more improvement. Moulton *et al*. showed that training junior surgical residents over a single multi-hour session versus training over multiple distributed sessions for a microvascular anastomosis skill resulted in both groups improving their performance. However, the distributed group performed significantly better on both the retention test and the transfer test[50]. A likely explanation for why the interval practice group benefits more than the mass practice group is due to the increased cognitive consolidation that is allowed to take place by creating time between training sessions[51]. Most current training programs that use proficiency criteria also put a limit on the amount of time devoted to practice during each session. The recommendation is typically to limit practice to 1 hour sessions[52]. What is less clear is the interval between practice schedules. A meta-analysis by Donovan and Radosevich indicated that simple tasks were better acquired over a shorter interval, whereas complex tasks required longer inter-training intervals[53].

Chapter 4. Enhancing Skill Retention

The goal of any training program is to convey long-term learning upon the learners. Long-term learning is best assessed by using retention tests. Retention is defined as the preservation of the aftereffects of experience and learning that makes recall or recognition possible[54]. It is generally believed that the ideal duration of the retention period depends on the skill of the learner and the training and testing environment. However, in order to get a true measure of learning, at least several days must pass to allow the temporary influences of practice to dissipate[18, 19]. Knowing how long a trainee retains the skills taught as part of a technical skills curriculum can greatly influence decisions on when and if any retraining schedules are required.

Repeated practice can produce temporary performance effects that disappear quickly after a session is finished or when the test condition changes[17, 55]. Many authors therefore consider retention tests conducted days to weeks after the initial practice session a better measure of true learning because they assess the durability of a response[18]. This more realistically reflects surgical residency, as it is rare to practice a task and then immediately perform it in a real life setting. More commonly, there is a delay before the opportunity to use a practiced skill presents itself in the operating room. Additionally, the ability to generalize or transfer a skill is another important measure of learning[55]. Since it is impossible to practice every individual skill under every possible condition, learners must be able to generalize competencies to different contexts. For surgical education then, the real value of simulator practice should not be measured by immediate post-practice performance, but rather by the ability to retain and generalize skills[18].

A tremendous amount of effort has been put into determining the optimal curriculum in order to maximize long-term learning and retention of skills. Plenty of original research and reviews in the motor learning literature have spent a considerable amount of time determining the moderators associated with skill decay. Skill decay refers to the loss of trained or acquired skills (or knowledge) after periods of nonuse. This area of research is particularly salient and problematic in situations where individuals receive formal training in a particular skill that they may not be required to use for extended periods of time. Two main categories of factors that influence the decay or retention of trained skills over extended periods of nonuse have been described: methodological and task-related factors[56]. Methodological factors are those that can be modified in the training or learning context to reduce skill loss. Examples of methodological factors include degree of overlearning, conditions of retrieval, and method of testing. Task-related factors, on the other hand, are inherent characteristics of the task and are typically not amenable to modification by the trainer, researcher, or both. Examples of task-related factors include physical vs. cognitive tasks and natural vs. artificial tasks. Since we cannot change the task-related factors of surgery and surgical skills, attention must focus on determining which methodological factors can be modified to allow for the least skill decay.

In a meta-analysis, Arthur *et al*. looked at skill retention[56]. They report on several factors that influence the decay or retention of trained skills over extended periods of nonuse. These factors are:

i. Length of the retention interval – Not surprisingly, a longer retention interval was associated with increased skill decay.

ii. The degree of overlearning – higher degrees of overlearning was associated with less skill decay over periods of nonuse.

iii. Task-specific characteristics – closed-loop tasks (those tasks with a definite beginning and end) resulted in more retention than open-looped tasks (continuous tasks with no definite beginning or end). Physical tasks resulted in less skill decay than cognitive tasks. Probably least surprising was the fact that speed tasks were more resistant to skill decay than accuracy tasks. In fact the decay for accuracy was more than three times that of speed.

iv. Conditions of retrieval – retention tests that were similar in context to the original learning environment resulted in greater retention.

v. Individual differences – motivation was shown to positively affect learning and retention.

While not all the factors that were addressed are directly related to complex motor skills, this meta-analysis contributes to several key concepts related to retention in the surgical skills literature.

A study of 16 first year surgical residents showed that skill retention was maintained at 7-8 months for the simpler FLS tasks of peg transfer and pattern cutting[57]. In contrast, the more difficult task of laparoscopic suturing did not show statistically significant skill retention at 7-8 months. The retention score for this study was only based on completion time however, so unfortunately no inference can be made on how time effects the amount of errors made by junior surgical residents.

Several authors have found that periodic retraining to proficiency standards reduces the deterioration of acquired technical skills over time. Castellvi *et al*. trained 42 surgery residents to proficiency in the FLS tasks and then had them undergo retention tests for both intracorporeal and extracorporeal suturing at 6.5 months and 12.5 months[58]. They found that for extracorporeal suturing, complete skill retention occurred in 45% of people at 6.5 months. Those not meeting proficiency standards underwent retraining and at 12.5 months 60% of trainees still met proficiency. For the intracorporeal suturing task, only 14% met proficiency at 6.5 months but with retraining for those who did not meet proficiency standards, 52% met proficiency at 12.5 months. The authors concluded that ongoing FLS training for laparoscopic suturing skills is beneficial and minimizes skill loss over time. Using the exact same initial study population, Mashaud *et al*. was able to show that after initial proficiency training, additional training with retention tests at 6 month intervals with mandatory retraining to proficiency if proficiency levels were not achieved, was successful at obtaining a 100% pass rate for the FLS certification examination[59].

Using a validated laparoscopic skills curriculum Stefanidis *et al.* trained 14 surgical residents to proficiency on both a virtual reality and video trainer simulator[60]. They found that after an average of 2 weeks, posttest scores deteriorated by 17%-45%, with a greater skill loss for the virtual reality trainer. However, retention tests occurring at an average of 7 months showed no further deterioration of laparoscopic skill. The initial amount of skill deterioration did not correlate with resident level or the amount of time taken to reach proficiency. This is further evidence that proficiency training has a durable effect on laparoscopic skills.

To further delineate how long laparoscopic suturing skills are retained for following completion of an expert-derived proficiency-based curriculum, Stefanidis et al. trained medical students to proficiency standards in intracorporeal suturing and then randomly assigned them to either ongoing retraining at 1 and 3 months, or no further training[61]. Overall they found that skill retention of more than 90% occurred by 6 months in both groups. Additionally there was no difference in scores at 2 weeks, 1 month, and 3 months. It was not until 6 months following training that a significant difference was seen between the ongoing training group and the control group. Ongoing training was able to decrease skill loss by 50% by adding only 2 retraining sessions within a 6-month period. Since no difference in skill loss for laparoscopic suturing was seen until 3 months, it would appear that the ideal time for initial retraining would occur somewhere between 3 and 6 months.

In contrast to the surgical literature that has only been studying skill retention for the better part of a decade, the non-surgical literature has been studying skill retention for several decades. Shadmehr and Brashers-Krug had participants practice specific patterns of arm movements in two distinct magnetic force fields using a blocked training schedule[62]. They found that the motor skills learned were able to be retained for at least 5 months. Interestingly, they also found that both tasks could be learned, but only if practice was separated by more than 5 hours. If the time interval between learning task 1 and task 2 is shorter, then the learning of the second task actually causes unlearning of the first task. This implies that the process of consolidating the learned material takes time, and that the representation of the recently acquired skill is fragile for a given period immediately following training. This has implication for surgical skills courses where several skills are being taught and practiced during a single session. Novel curricula are required to ensure that the trainee gets the most learning out of the training sessions.

Motor skills can be acquired and retained in two different forms, accuracy and speed. Hikosaka *et al*. had subjects learn a sequence of button presses over an 8-10 day learning session[63]. Retention tests were held 16 months later where they found that both speed and accuracy were retained. However, speed was retained to a greater extent than accuracy was. This shows that not all aspects of performance degrade at the same rate.

For programs or courses that consists of multiple discrete tasks, practice can be scheduled in a blocked or random fashion. A training protocol, in which a single task is practiced repeatedly before moving on to the next task, is an example of blocked practice. In contrast, random practice schedules involve performing the same number of repetitions of each task, but in a random order such that a given task is never practiced on successive trials. This varying of practice schedules, or other learning variables, has been termed “contextual interference” (CI) in the motor learning literature[64]. Contextual interference occurs in two forms: high and low. Contextual interference is termed “high” when at least one variable in a given task or practice schedule is altered. Alternatively, contextual interference is termed “low” when there is minimal or no variation from task to task. An example of high contextual interference is arranging the practice schedule into random orders. An example of low contextual interference is continually practicing the same golf shot with the same golf club over and over before moving on to the next shot. The contextual interference effect suggests that high contextual interference (random practice schedules) allows for greater retention and transfer of motor skills, but that low contextual interference (blocked practice schedules) allows for faster acquisition of motor skills[15]. It has been suggested that manipulating certain variables, such as practice schedules, while learners are in an acquisition phase in order to slow down the speed of acquisition may actually increase long-term learning and transfer of motor skills. Conversely, conditions that speed the rate of improvement may actually be detrimental to long-term learning and transfer of skill[55].

In the motor literature, Shea and Morgan[15] first brought forward the idea that training schedules appears to be an important variable in modifying the effect of skill retention. In a study that required participants to knockdown barriers in a predefined sequence, with each sequence representing a different task, participants performed immediate post-tests following training and then delayed retention tests (given under either random or blocked conditions) ten days later. On the immediate post-tests there was a clear advantage found for the blocked practice group. However, on the delayed retention tests, there was an advantage for the random practice group regardless of whether the retention tests were given in a random or blocked order. Repeated practice in a blocked fashion may help participants to acquire motor skills more quickly, but they do not necessarily retain these skills over time. Conversely, participants who follow random practice schedules may acquire skills more slowly, but they better retain their skills over time. These findings cannot be explained by learning specificity, which postulates that the degree of overlap between training and test conditions determines test performance, since the random schedules conferred an advantage regardless of whether the retention tests were given in a blocked or random order. Instead, the authors hypothesized that the random group was forced to use multiple processing strategies during the acquisition phase, whereas the blocked group were not required to use such multiple processing strategies[15]. Whether this phenomenon exists within the surgical domain has received little study. Current surgical training programs commonly use blocked practice schedules, with the assumption that the conditions that improve the speed of skill acquisition will also support long-term learning.

Chapter 5. Objectives

An understanding of how practice schedules influence learning has broad implications for laboratory-based technical skills courses and curricula. Current training programs commonly use blocked practice schedules, with the assumption that the conditions that improve the speed of skill acquisition will also support long-term learning. The FLS course is perhaps the most notable example of this type of practice. It is performed by thousands of trainees each year as the standard for learning basic laparoscopic skills in North America[8]. Findings of an advantage for blocked practice regimens would support current models of technical skills education and require reconsideration of motor learning principles in light of complex surgical tasks. However, findings of advantage for random practice schedules should cause wide and critical re-examination of current programs to ensure they are designed to optimize long-term learning and skill transfer.

In the motor learning literature, authors have examined the effects of altering schedules of task practice. Their findings suggest that adding variability to practice schedules may enhance long-term learning. However, surgery is more complex than other motor skills and whether these findings can be applied to surgical skills education is unclear.

The objectives of this study are:

1. To determine the optimal schedule of task practice for learning basic laparoscopic skills

a. To assess the effects of blocked and random schedules of task practice on the acquisition of laparoscopic skills
b. To assess the effects of blocked and random schedules of task practice on the retention and transfer of laparoscopic skills

Chapter 6. Methods

A. Participants

General surgery, urology, plastic surgery, orthopedic surgery, otolaryngology, cardiac surgery, and obstetrics and gynecology residents in postgraduate years 1-2 (PGY – 1-2) at the University of Manitoba were invited to participate in this study. Senior medical students with a self-declared interest in a surgical specialty attending the University of Manitoba were also invited to participate. Learning basic laparoscopic surgical techniques is a fundamental part of residency training for general surgery, urology, and obstetrics and gynecology. Plastic surgery, orthopedic surgery, otolaryngology, and cardiac surgery do not get formal laparoscopic training. However, they all get formal training in other minimally invasive and microscopic techniques. Medical students and residents were stratified separately and randomization occurred using sealed envelopes.

Since we are interested in laparoscopic skill acquisition, residents who had completed the FLS course or had significant experience in laparoscopic surgery (defined as the primary surgeon for any laparoscopic procedure) were excluded from participation. All potential participants completed a questionnaire to determine eligibility prior to entry into the study (Appendix 1). In addition to the questionnaire, informed consent was obtained from all participating subjects (See Appendix 2 for a copy of the informed consent form).

Ethics approval was obtained by the University of Manitoba Health Research Ethics Board (HREB).

B. Power Calculation and Estimated Sample Size

Similar research in skills training and transfer has demonstrated large effect sizes of 1.8 standard deviations (SD) when comparing various treatment and control groups[65]. To detect an effect size of 1.8 SD, using a 2-tailed alpha of 0.05 and a power of 0.8, we estimated that 7 subjects would be required in the control group. Given the anticipated attrition due to residents' schedules, we recruited 9 participants into the control arm. In the field of psychology an effect size of 1 SD is considered a large but acceptable difference in assessing teaching intervention[66]. Additionally, several other studies have used effect sizes between 1 and 1.2 in order to calculate sample sizes[65, 67]. The variability depends on the amount of practice and participant skill and training level. As this study is similar in design to previous studies comparing different treatment groups, we expected to obtain a similar effect size. To detect an effect size of 1.2 SD, using a 2-tailed alpha of 0.05 and a power of 0.8, we estimated that 12 subjects would be required in each practice schedule arm.

C. Study Design

We used a randomized, controlled study design, incorporating multiple outcome measures. Data collection occurred between October 2011 and March

2012. After reviewing the consent form and filling out the questionnaire, all participants were oriented to the FLS box-trainer. Following this, they then watched a video tutorial demonstrating each task (FLS CD ROM Disk #2). The video explained the objectives of the tasks while simultaneously showing a live demonstration, emphasizing the key components. A real-world example of how that particular skill can be incorporated into an actual surgical procedure was then provided.

Participants then performed each task twice. The average of these two trials were then taken and this constituted their baseline pre-test scores. After completion of the pre-test all individuals were then randomized into one of three groups using unmarked, sealed envelopes: (1) no additional training (control group), (2) FLS training using a blocked task schedule, or (3) FLS training using a random task schedule. Medical students and residents were stratified separately. Figure 1 shows a flow diagram of the study design.

The manual skills portion of FLS consists of five tasks[10]:

1. **Peg transfer**

   The operator uses laparoscopic graspers to lift a peg from one peg board with the left hand, transfer it the right hand, and then place it on a second peg board. This sequence is repeated with five more pegs. The entire exercise is then repeated in reverse. The cut-off time is 300 seconds. A penalty score is given by calculating the percentage of pegs that could not be transferred as a result of being dropped outside of the field of view.

2. **Pattern cutting**

The operator uses laparoscopic scissors to cut a pre-marked circular pattern out of a 10 x 10 cm suspended sheet. The grasper in the non-dominant hand is used to place the material under tension in order to facilitate the task. The cut-off time is 300 seconds. The penalty is determined by calculating the percentage of the area of deviation from the original pattern. For simplicity we added a standard penalty of 75 points for any deviation from the pattern by more than 2mm.

3. **Ligating loop application**

The operator places a pre-tied slipknot at a specific pre-marked point on a foam tubular appendage. It requires the operator to stabilize the appendage, place the loop on the target, secure the knot, and cut the excess suture. The cut-off time is 180 seconds. The penalty is the distance in millimetres that the knot deviates from the pre-marked point determines the penalty score. An additional 50-point penalty is applied for insecure or failed knots.

4. **Intracorporeal (IC) suturing**

The operator places a single simple stitch (2-0 silk suture 12 cm in length on a curved v-20 needle) through pre-marked points on a longitudinally slit Penrose drain. This task requires the operator to transfer and place the needle and perform an instrument tie. Three throws are then placed to tie the suture using an intracorporeal technique. The first knot is a double throw, the second and third knots are single throws. The cut-off time is 600 seconds. The penalty score is the sum of the distance in millimetres that the suture placement deviates from the pre-marked points plus the gap in millimetres between the edges of the approximated tissue. An additional penalty is applied for knots that slip (10 points) or come apart (20 points).

5. **Extracorporeal (EC) suturin**g

   This is identical to the previous task, except that a longer suture is used, and the knot is tied using an extracorporeal technique that uses a knot pusher. The cut-off time is 420 seconds, with a penalty as assessed in task 4.

We excluded the EC suturing task in order to reduce the overall duration of practice for study purposes. Thus, this study included tasks 1-4 from the above list. Participants randomized to either training group performed the following number of repetitions for each task in their practice sessions: 12 peg transfers, 6 pattern cuts, 4 ligating loop applications, and 10 IC suturing. Previous studies suggest that participants will show measurable improvements in task performance with these repetition numbers, although they are very unlikely to achieve expert levels of proficiency[12, 14, 45, 68]. Participants in the blocked schedule group performed all repetitions of a single task before moving on to the next task. Participants in the random schedule group performed the same number of repetitions of each task but in a random order, with no task being repeated on two consecutive trials. The total number of repetitions was constant between experimental arms, only the order of task practice varied. Figure 2 shows the practice schedule used for all three groups.

Figure 1. Flow diagram of study design

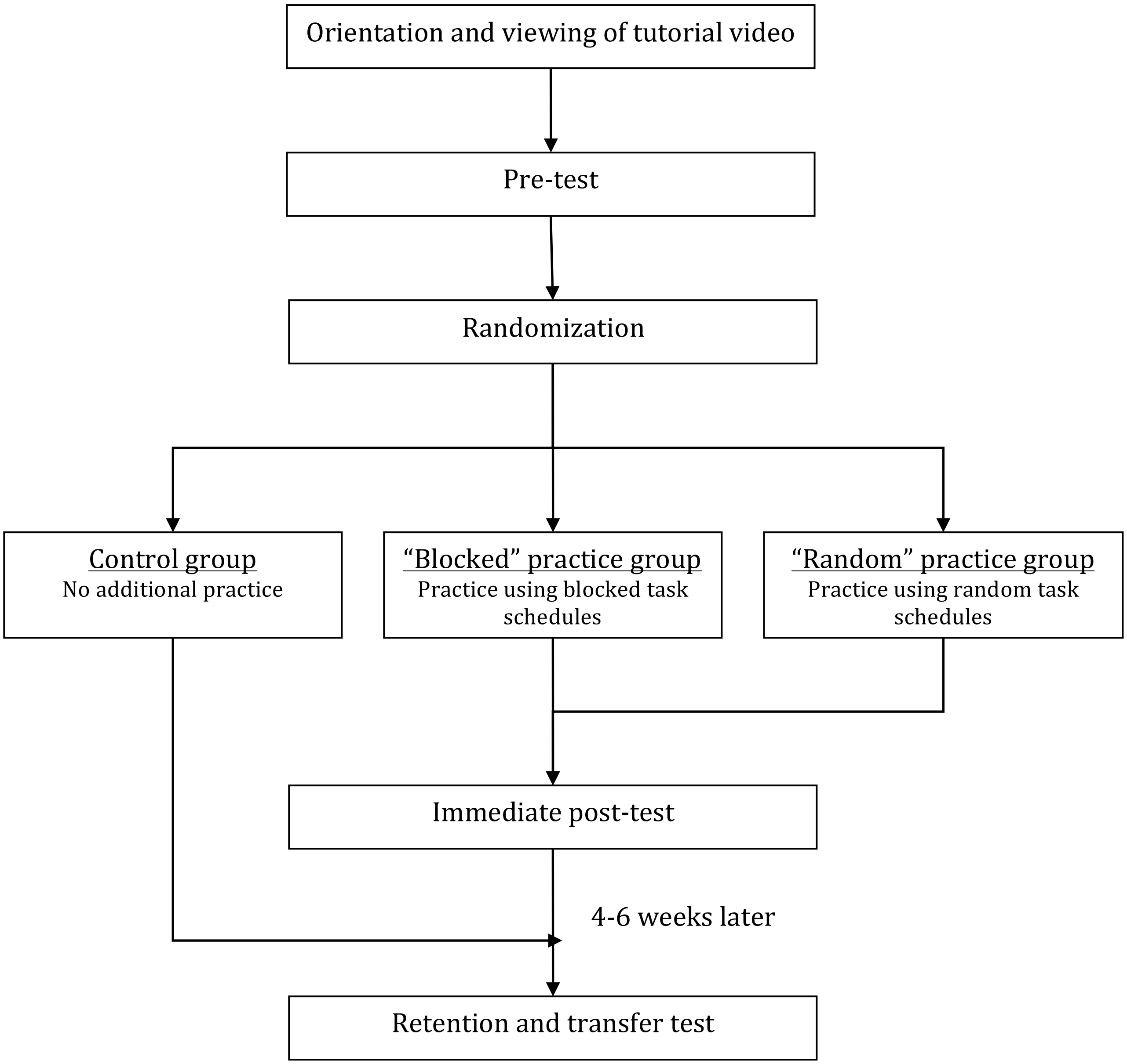


Figure 2. Practice schedule used for the three groups*

	Group 1 (No Additional Practice)	Group 2 (Blocked Task Schedule)	Group 3 (Random Task Schedule)
Task		Peg Transfer	Peg Transfer
		Peg Transfer	IC Suturing
		Peg Transfer	Pattern Cut
		Peg Transfer	Peg Transfer
		Peg Transfer	IC Suturing
		Peg Transfer	Peg Transfer
		Peg Transfer	Ligating Loop
		Peg Transfer	IC Suturing
		Peg Transfer	Peg Transfer
		Peg Transfer	Pattern Cut
		Peg Transfer	Peg Transfer
		Peg Transfer	IC Suturing
		Pattern Cut	Pattern Cut
		Pattern Cut	IC Suturing
		Pattern Cut	Ligating Loop
		Pattern Cut	Peg Transfer
		Pattern Cut	IC Suturing
		Pattern Cut	Pattern Cut
		Ligating Loop	Peg Transfer
		Ligating Loop	IC Suturing
		Ligating Loop	Peg Transfer
		Ligating Loop	Pattern Cut
		IC Suturing	IC Suturing
		IC Suturing	Ligating Loop
		IC Suturing	Peg Transfer
		IC Suturing	IC Suturing
		IC Suturing	Ligating Loop
		IC Suturing	Peg Transfer
		IC Suturing	Pattern Cut
		IC Suturing	Peg Transfer
		IC Suturing	IC Suturing
		IC Suturing	Peg Transfer

*IC, Intracorporeal

All participants in the blocked and random practice groups performed a post-test immediately after training as a measure of skill acquisition. The final trial of each task constituted the post-test. All three groups then performed retention and transfer tests 4-6 weeks after training. The retention tests consisted of performing two sets of the four tasks, once in the blocked order and once in the random order. Half the participants in each experimental group performed the test tasks first in a blocked and then in a random order, while the other half performed the test tasks first in a random followed by a blocked order. The transfer of skill test consisted of a novel, unpracticed laparoscopic cannulation task[8, 12, 69], which was performed twice in succession. The cannulation task involves suspending intravenous tubing with a premade defect in the endotrainer box. After viewing an instructional video, the participants use 2 curved forceps to introduce a cholangiogram catheter into the box, thread it into the tubing up to a marked length (4 cm) and then remove it. Subtracting the time taken to complete the task from a cutoff time of 300 seconds creates the score. The score is then normalized against the best overall score from the sample.

D. Outcome Measures

Outcome measures on each of the tests included: (1) FLS scoring system metrics[10, 12, 70] and (2) computer-based evaluations[38, 71-75]. Both forms of outcome measures have been previously validated. The transfer task has recently been validated and shown to have excellent concurrent validity with the other FLS tasks[76]. In the FLS scoring system, a cut-off time is assigned for each task[10, 12]. A raw score was then calculated by subtracting the time taken to complete the task (in seconds) and any additional penalty points from the cut-off time, as follows:

Raw score = Cut-off time – Completion time – Penalty score

Consequently, higher scores indicated superior performance. Raw scores were then normalized to provide equal weighting of the tasks, according to formulas available through the FLS program[68]. By this scoring system, a score of 100 on any individual task indicates excellent performance. For the pre-, post-, and retention tests, the normalized score on each of the four individual tasks were summed to give an overall test score.

The computer-based performance measures consist of completion time and hand motion analysis (number of hand movements and path-length travelled). Hand motion data were collected with an electromagnetic (EM) tracking system (Ascension Technology Corporation, Burlington, VT, USA) and analyzed using a customized software program (MARS: Motion Analysis and Recording System – University of Manitoba). This system tracks and records the position of small sensors (6 degrees of freedom) placed on the dorsum of each of the participants' hands. The software translates the raw three-dimensional positional data from the x,y,z Cartesian coordinates into scores of dexterity, including number of movements, path length travelled, and time. For motion tracking purposes, a movement is defined as a change in velocity. MARS uses a Gaussian filter with a width of 12 samples per standard deviation, which corresponds to a high frequency cutoff filter of 1.666Hz when the sample rate is 20Hz. This allows us to smooth out the velocity curves and eliminate the unwanted high frequency background noise (such as hand tremors), ensuring that only meaningful actions are recorded and reported. Because a low pass filter, like the Gaussian filter, will never eliminate all the unwanted noise, we also instituted a velocity threshold above which the software will report an actual movement. We set the velocity threshold at 15 mm/second. These settings were chosen based on previous calibration experiments reported in the literature[35]. The MARS software reports the number of movements, total distance travelled, and time taken to complete the task. Since this motion analysis is automated, the trainee can receive feedback without direct human supervision. The data collected from the left and right hands were added together to give a total number of movements and path-length travelled for each task. Some of the many benefits of hand motion analysis systems are that they are portable and have broad utility in their ability to be used in laparoscopic, endoscopic, and open procedures. Hand motion analysis as a measure of technical skill has previously been validated for open and laparoscopic tasks[38, 71-73, 75].

For both the FLS scores and the computer-based performance measures, subtracting the pre-test scores from the post-test scores allowed us to measure skill acquisition. Higher scores indicated a higher level of skill acquisition. Additionally, skill retention was calculated by subtracting the pre-test score from the retention score. Again, higher scores indicated more skill retention.

E. Data Analysis

The data was initially assessed for normality using the Shapiro-Wilks test and by inspecting the histograms. Homogeneity of variance was assessed using the Levene's test. This was performed for each of the outcome measures. This showed that for the FLS scores, the data was normally distributed and therefore allowed us to use parametric tests. Separate one-way analysis of variance (ANOVAs) and one-way analysis of covariance (ANCOVAs) with groups as the factor and pretest scores as the covariate were used to analyze all the FLS normalized task scores as well as the normalized overall scores, acquisition scores, and retention scores. Paired t-tests were used to assess the amount of skill acquisition and retention that occurred within each group. Most of the variables in the computer-based hand-motion data (time, path length, and number of movements) were assessed to be in a non-normal distribution, and therefore nonparametric tests were used to analyze this portion of the data. Kruskal-Wallis tests were used to compare the three groups. Significant differences at $P < 0.05$ were further assessed using the Mann-Whitney U test with bonferroni correction, giving us a significance level of $P < 0.017$. Paired data were assessed using the Wilcoxon signed ranks test. Data were analyzed using the SPSS® software, version 17 (Chicago, Illinois, USA).

Chapter 7. Results

A. Participant Characteristics

A total of 36 participants (14 male, 22 female) volunteered to participate in the study (See Table 1 for the baseline characteristics of all participants). Participants included senior medical students (n = 10), postgraduate training year (PGY) 1 (n = 17), and PGY 2 (n = 9). No participants were excluded from the study on the basis of extensive prior laparoscopic experience. PGY specialties included general surgery (n = 5), urology (n = 2), plastic surgery (n = 2), otolaryngology (n = 2), obstetrics and gynecology (n = 11), orthopedic surgery (n = 3), and cardiac surgery (n = 1). There was not a significant difference between the three study groups' characteristics at baseline. Thirty-two participants were right-handed and 4 were left-handed. The blocked group contained 13, the random group contained 14, and the control group contained 9 participants. All 36 participants completed the study. The average time between the training sessions and the posttest and retention test was not significantly different between the three groups.

Table 1. Characteristics of Study Participants

	Blocked (N=13)	Random (N=14)	Control (N=9)
Age (SD)	28 (3.6)	29 (4.1)	29 (5.7)
Gender (M:F)	6:7	5:9	3:6
Handedness (L:R)	1:12	3:11	0:9
Experience			
PGY-1	8 (62%)	7 (50%)	2 (22%)
PGY-2	2 (15%)	4 (29%)	3 (33%)
Medical Student	3 (23%)	3 (21%)	4 (44%)
Home Program			
General Surgery	3 (23%)	2 (14%)	0 (0%)
Plastic Surgery	2 (15%)	0 (0%)	0 (0%)
Urology	0 (0%)	2 (14%)	0 (0%)
Obstetrics & Gynecology	2 (15%)	6 (43%)	3 (33%)
Otolaryngology	1 (8%)	0 (0%)	1 (11%)
Orthopedic Surgery	1 (8%)	1 (7%)	1 (11%)
Cardiac Surgery	1 (8%)	0 (0%)	0 (0%)
Medical Student	3 (23%)	3 (21%)	4 (44%)
Time between training and retention test in days (SD)	38.2 (7.9)	39.4 (7.3)	39.8 (7.3)

The participant characteristics were generally similar between the groups. The age difference between the groups was not significantly different ($F[2,33] = 0.40$, $p = 0.68$). Similarly, there was not a significant difference seen in gender ($\chi^2[2] = 0.47$, $p = 0.79$), handedness ($\chi^2[2] = 2.79$, $p = 0.25$), experience ($\chi^2[4] = 3.79$, $p = 0.436$), and program ($\chi^2[14] = 18.28$, $p = 0.308$). The amount of time between the initial pretest and training and the retention tests was also not significantly difference between the groups (F[2,33] = 0.131, p = 0.88).

B. FLS Scores

1) Pretest

Pretest FLS scores were not significantly different between the three groups for the 4 individual tasks or for the overall combined score (see Table 2).

Table 2. Pretest FLS mean scores (SD) for the 3 study groups

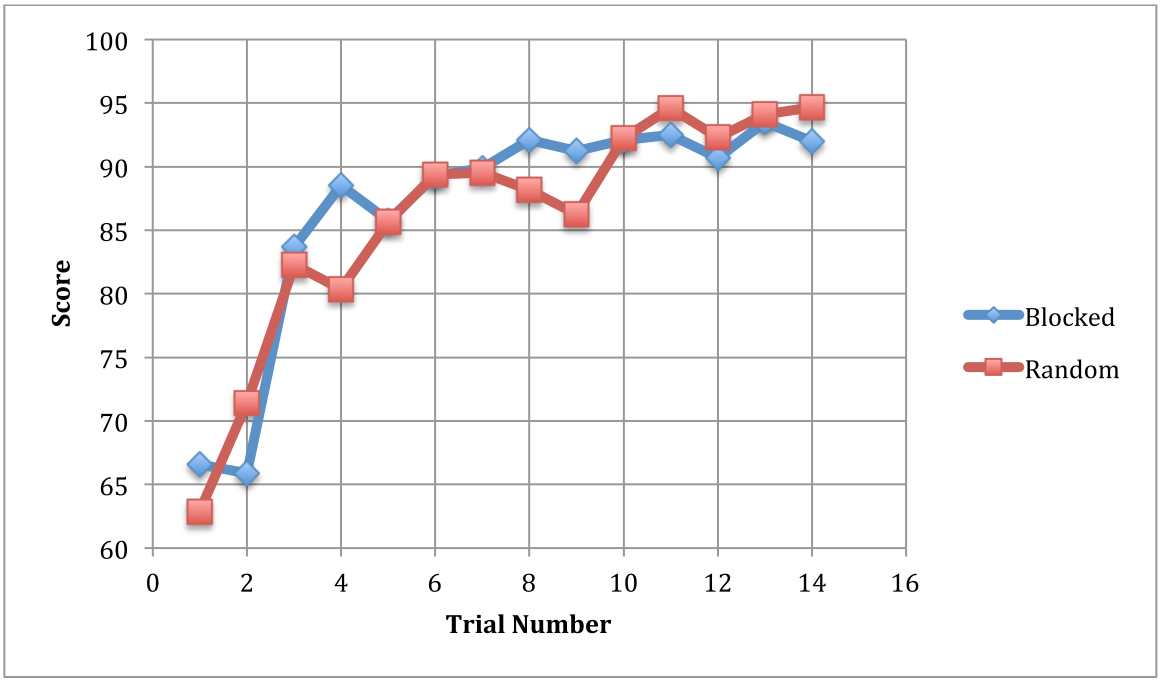
	Blocked	Random	Control	Significance Level
Peg Transfer	66.2 (19.3)	67.1 (16.3)	71.5 (11.3)	F[2,33] = 0.30, p = 0.74
Pattern Cutting	16.3 (15.1)	11.2 (14.3)	18.9 (17.7)	F[2,33] = 0.75, p = 0.48
Ligating Loop	42.7 (18.5)	47.3 (14.1)	50.6 (13.4)	F[2,33] = 0.70, p = 0.51
IC Suturing	35.6 (16.9)	33.0 (22.5)	44.6 (17.9)	F[2,33] = 1.02, p = 0.37
Overall Score	160.9 (47.9)	158.7 (48.9)	185.6 (38.7)	F[2,33] = 1.07, p = 0.36

* IC = Intracorporeal, SD = Standard Deviation

2) Performance Curves

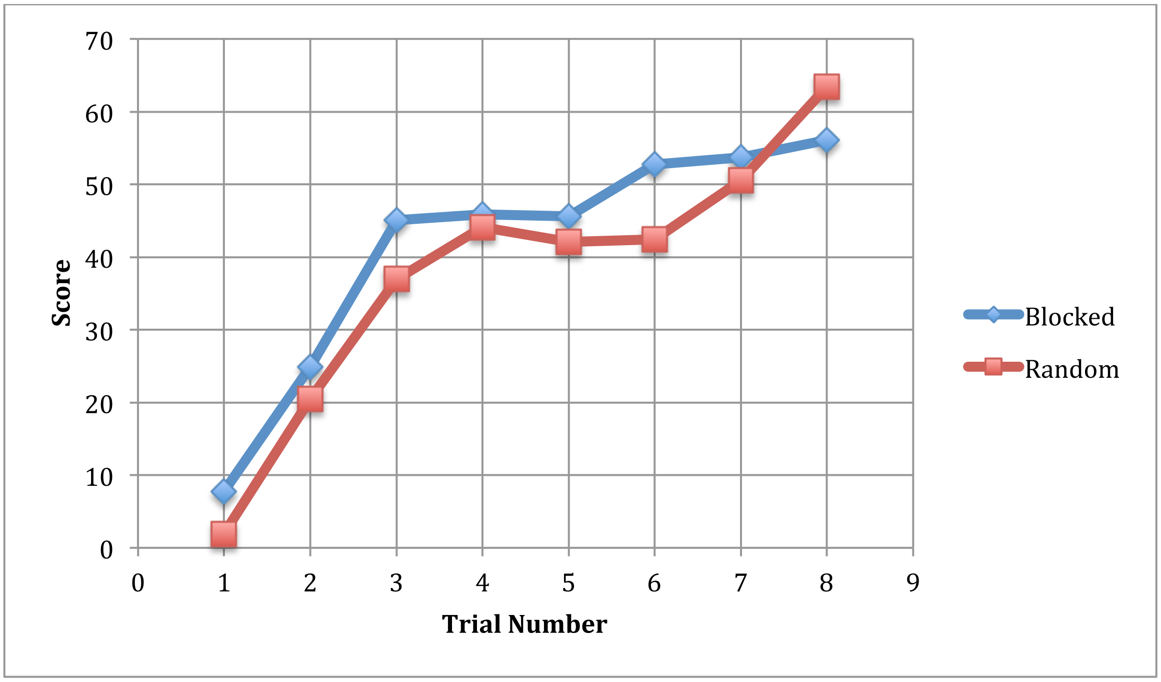
Upon completion of the training schedules, both the blocked group and the random group showed statistically significant improvement from their pretest scores to their posttest scores. Figures 3 through 6 show the performance curves generated for both the blocked and the random groups for all 4 tasks. There was also a statistically significant improvement in the overall score when comparing the pretest to the posttest score in both the blocked (mean improvement score = 137.9, SD = 39.5, t[12] = 12.60, p<0.0005; paired t-test) and the random group (mean improvement score = 144.4, SD = 52.4, t[13] = 10.31, p<0.0005; paired t-test).

Figure 3. Performance curve for the Peg Transfer task FLS scores



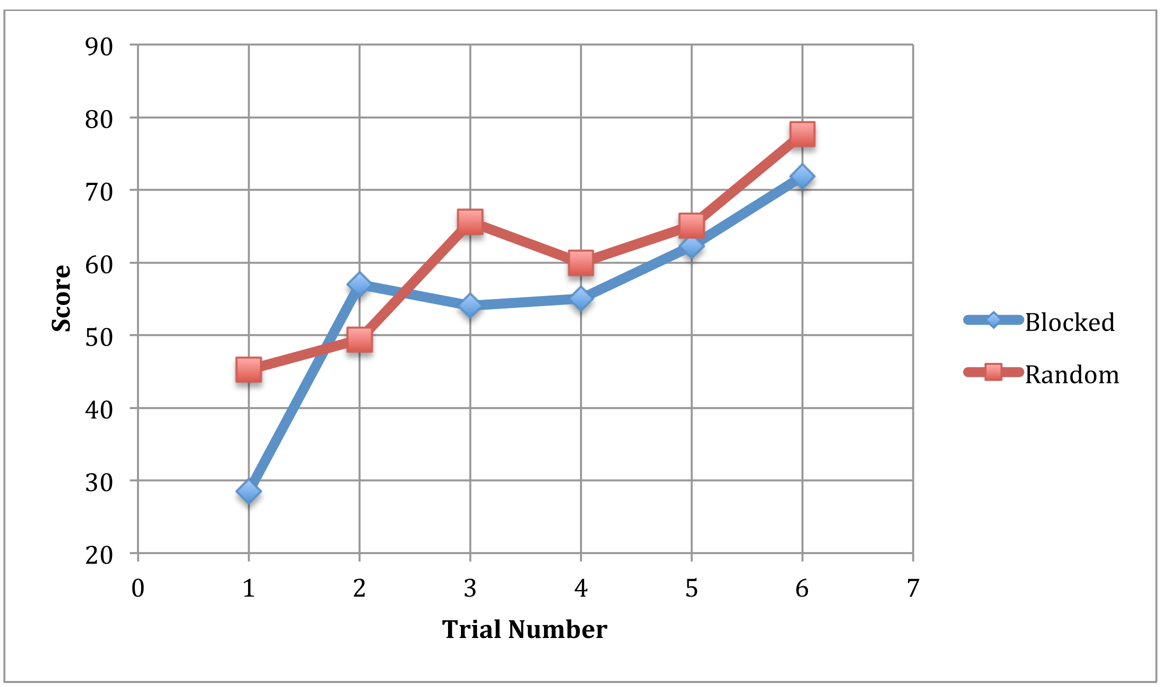
Blocked Post vs. Pretest: mean improvement = 25.8, SD = 18.2, t[12] = 5.11, p <0.0005
Random Post vs. Pretest: mean improvement = 27.5, SD = 13.2, t[13] = 7.79, p <0.0005

Figure 4. Performance curve for the Pattern Cutting task FLS scores



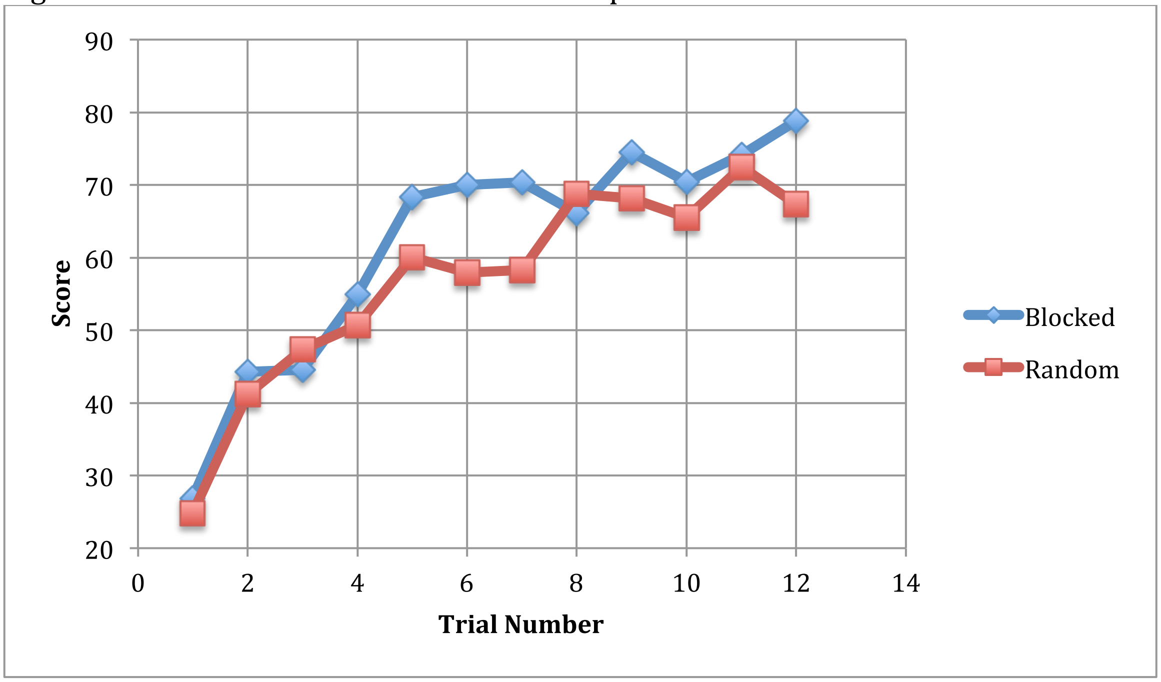
Blocked Post vs. Pretest: mean improvement = 39.8, SD = 21.5, t[12] = 6.68, $p<0.0005$
Random Post vs. Pretest: mean improvement = 52.2, SD = 16.4, t[13] = 11.94, $p<0.0005$

Figure 5. Performance curve for the Ligating Loop task FLS scores



Blocked Post vs. Pretest: mean improvement = 29.1, SD = 20.9, t[12] = 5.03, $p<0.0005$
Random Post vs. Pretest: mean improvement = 30.4, SD = 18.9, t[13] = 6.02, $p<0.0005$

Figure 6. Performance curve for the Intracorporeal Suture task FLS scores



Blocked Post vs. Pretest: mean improvement = 43.3, SD = 16.0, t[12] = 9.75, $p<0.0005$
Random Post vs. Pretest: mean improvement = 34.4, SD = 27.2, t[13] = 4.74, $p<0.0005$

3) Posttest

Immediately following training, the blocked group and the random group performed each of the 4 tasks as a posttest. There was not a significant difference seen between the blocked and random groups for either the individual tasks or the overall score (see Table 3).

Table 3. Posttest FLS mean scores (SD) for blocked and random groups

	Blocked	Random	Significance Level
Peg Transfer	92.0 (7.3)	94.7 (5.3)	t[25] = 1.10, p = 0.28
Pattern Cutting	56.1 (23.8)	63.4 (18.2)	t[25] = 0.90, p =0.38
Ligating Loop	71.9 (17.2)	77.7 (12.2)	t[25] = 1.01, p = 0.32
IC Suturing	78.8 (8.3)	67.4 (24.4)	t[25] = 1.61, p = 0.12
Overall Score	298.8 (35.3)	303.1 (31.2)	t[25] = 0.33, p = 0.74

* IC = Intracorporeal, SD = Standard Deviation

Improvements in performance from the pretest to the posttest scores for the blocked group and the random group were assessed using a one-way analysis of covariance (ANCOVA), using the pretest scores as the covariate. The results were not significantly different between the blocked and random groups (see Figures 7-11) for the Peg transfer (F[1,24] = 1.34, p = 0.26), the Pattern cut (F[1,24] = 2.16, p = 0.15), Ligating loop (F[1,24] = 0.75, p = 0.40), Intracorporeal suturing (F[1,24] = 2.41, p = 0.13), and the overall scores (F[1,24] = 0.16, p = 0.69).

Figure 7. Improvement in Peg Transfer score after training

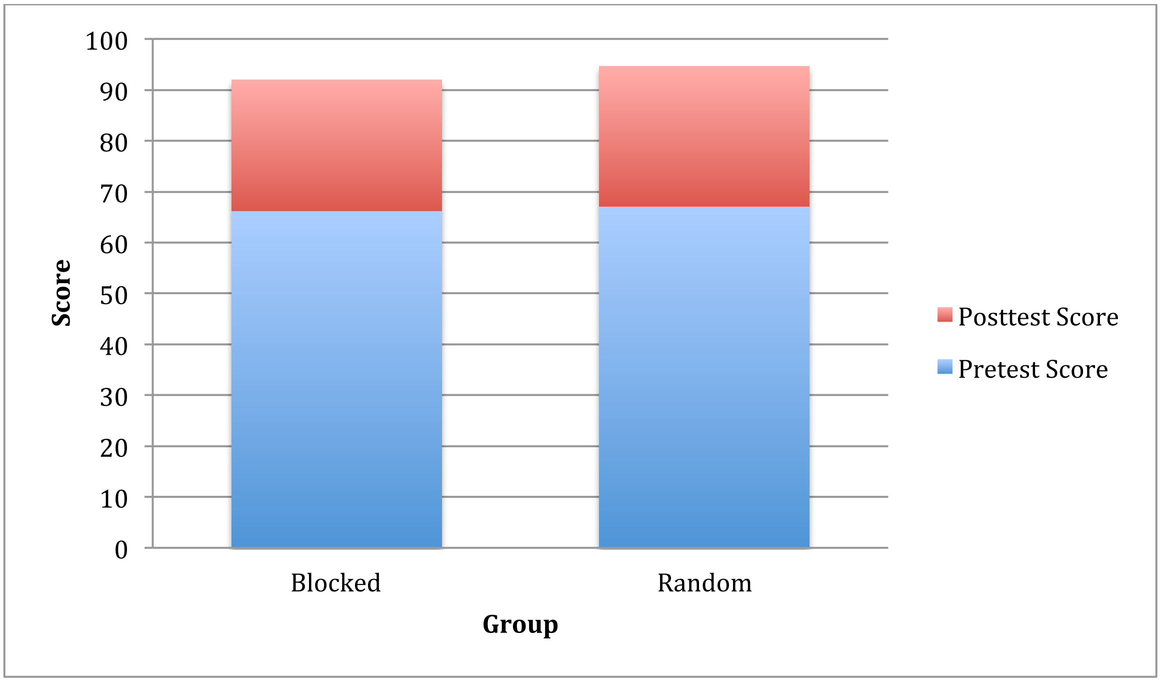


Figure 8. Improvement in Pattern Cutting score after training

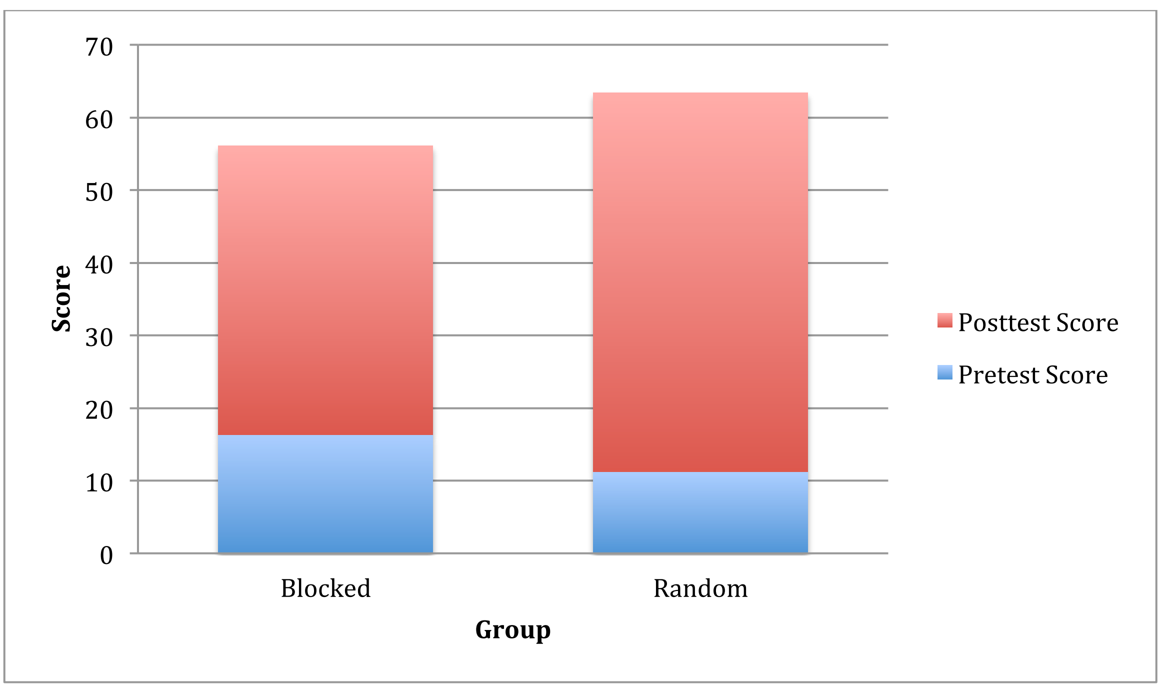


Figure 9. Improvement in Ligating Loop score after training

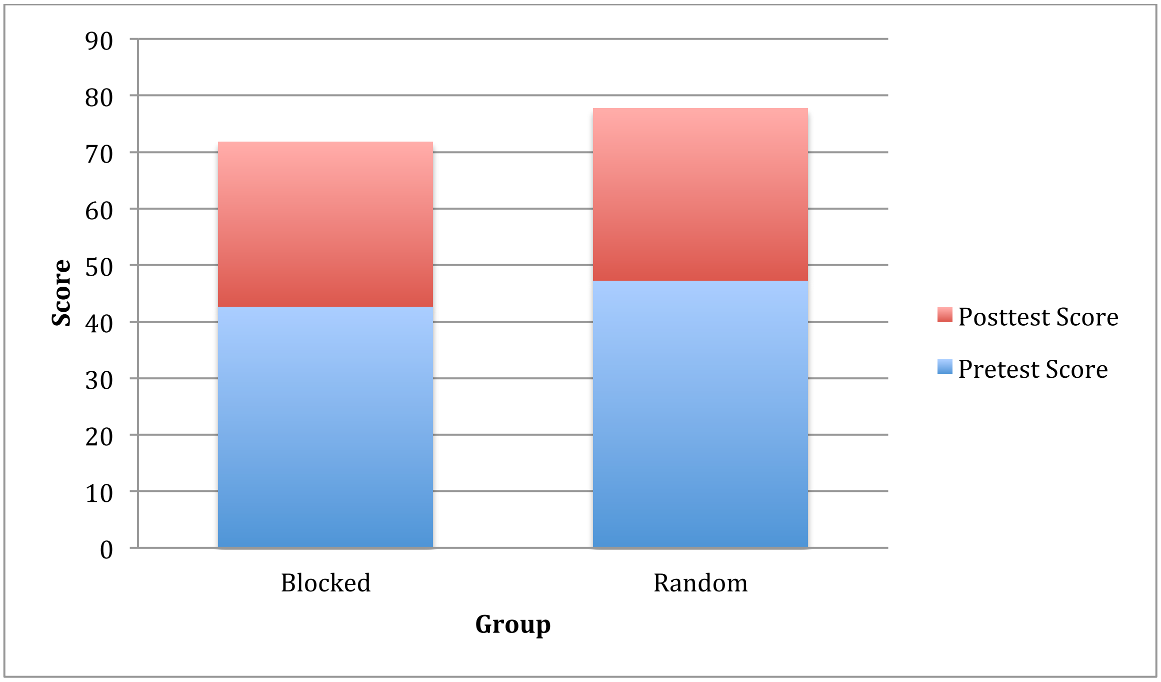


Figure 10. Improvement in Intracorporeal Suturing score after training

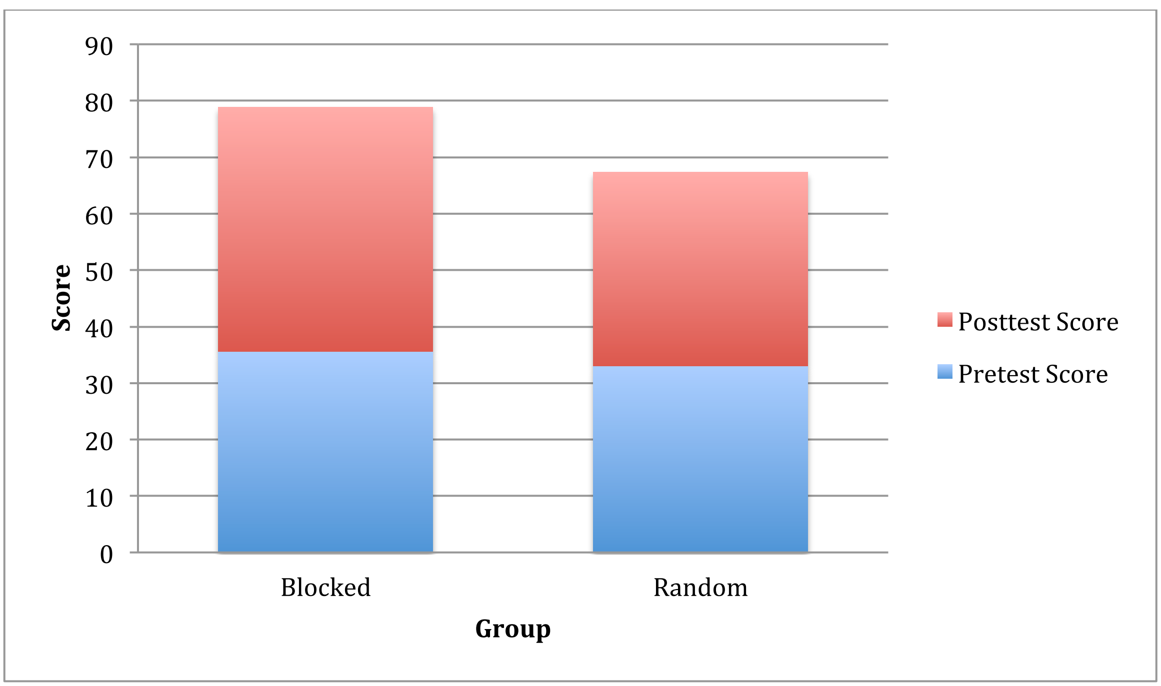
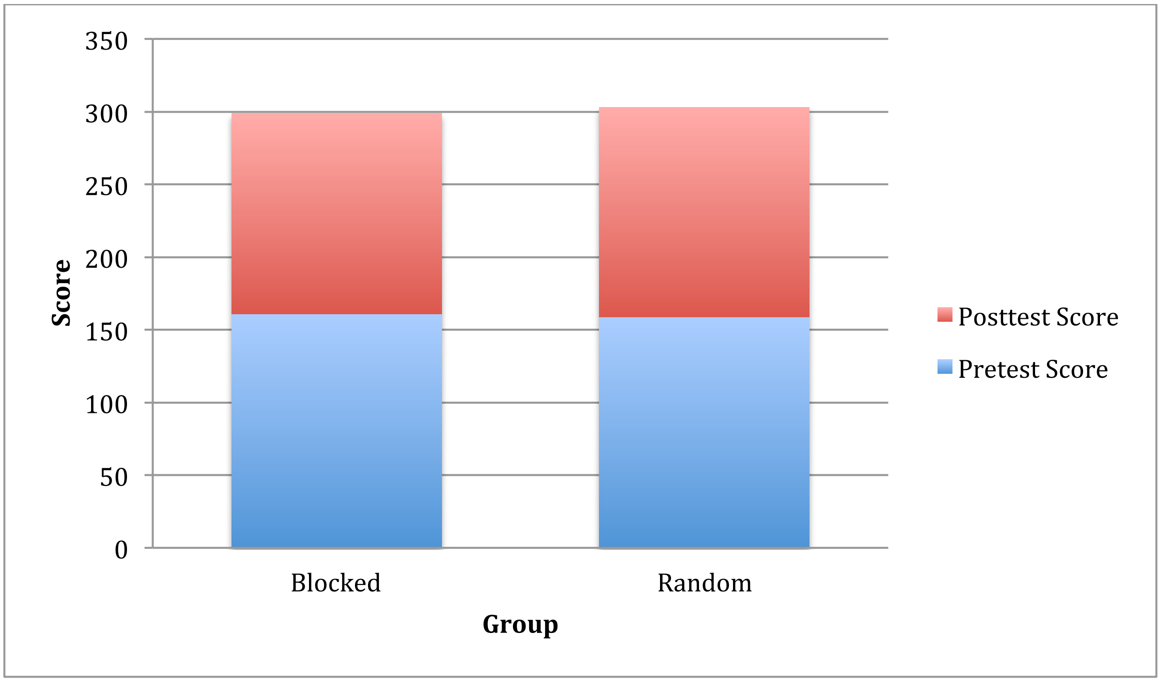


Figure 11. Improvement in Overall score after training



4) Skill Retention

All three groups returned to perform retention tests 4-6 weeks following the initial training (see Table 1). All participants performed each of the 4 tasks as their retention test. Table 4 shows the paired t-test results comparing the posttest and the retention test scores for the blocked group. Three of the 4 tasks showed a statistically significant decline in performance. Table 5 shows the paired t-test results comparing the posttest and the retention test scores for the random group. Again, 3 of the 4 tasks showed a significant decline in performance.

Table 4. Comparison of retention and posttest FLS mean scores (SD) for the blocked group

	Posttest score	Retention score	Significance Level
Peg Transfer	92.0 (7.3)	87.2 (6.0)	t[12] = 2.27, p = 0.04
Pattern Cutting	56.1 (23.8)	40.1 (29.0)	t[12] = 1.69, p = 0.12
Ligating Loop	71.9 (17.2)	55.9 (15.6)	t[12] = 2.62, p = 0.02
IC Suturing	78.8 (8.3)	60.1 (23.2)	t[12] = 3.24, p = 0.01
Overall Score	298.8 (35.3)	243.3 (45.6)	t[12] = 4.11, p = 0.001

IC = Intracorporeal, SD = Standard Deviation

Table 5. Comparison of retention and posttest FLS mean scores (SD) for the random group

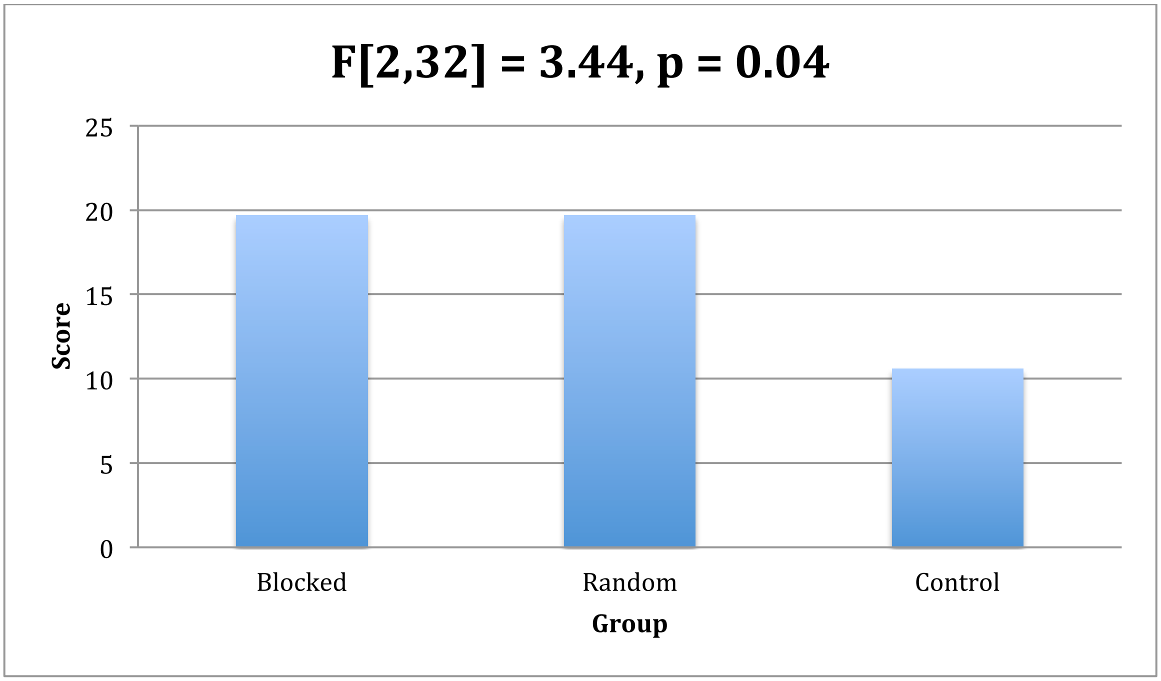
	Posttest score	Retention score	Significance Level
Peg Transfer	94.7 (5.3)	87.4 (12.7)	t[13] = 2.97, p = 0.01
Pattern Cutting	63.4 (18.2)	34.5 (24.9)	t[13] = 3.23, p = 0.01
Ligating Loop	77.7 (12.2)	44.5 (25.3)	t[13] = 4.23, p = 0.001
IC Suturing	67.4 (24.4)	59.9 (28.8)	t[13] = 0.87, p = 0.40
Overall Score	303.1 (31.2)	226.4 (68.7)	t[13] = 4.14, p = 0.001

IC = Intracorporeal, SD = Standard Deviation

5) Skill Learning

The amount of skill learning (retention score subtract pretest score) for each of the 4 tasks and the overall score are depicted in Figures 12-16. Again an ANCOVA was conducted with the pretest score as the covariate. Post-hoc Tukey's tests results showing where the statistically significant difference lies are also depicted. Overall, scores for both the blocked and random groups were better than for the control group. Where significant differences existed, the difference was between the practice groups (blocked and random) and the control group. No difference was seen between the blocked and the random group.

Figure 12. Skill learning in the Peg Transfer task, based on FLS score



*Post-hoc Tukey's comparisons show no significant differences

Figure 13. Skill learning in the Pattern Cutting task, based on FLS score

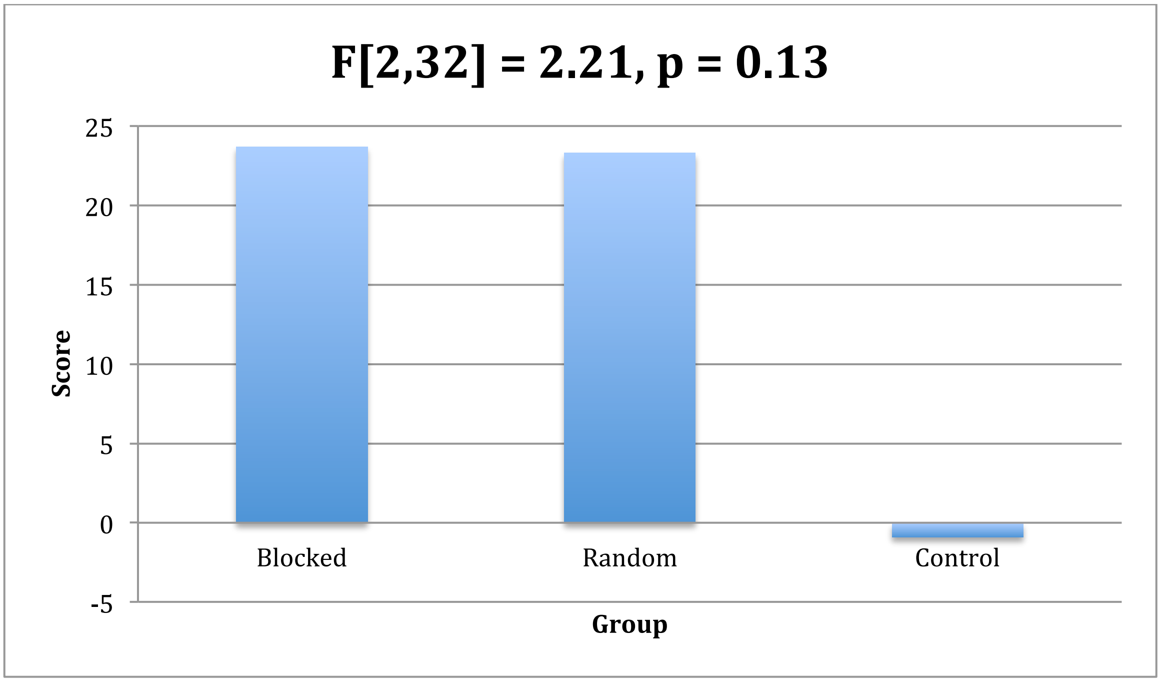
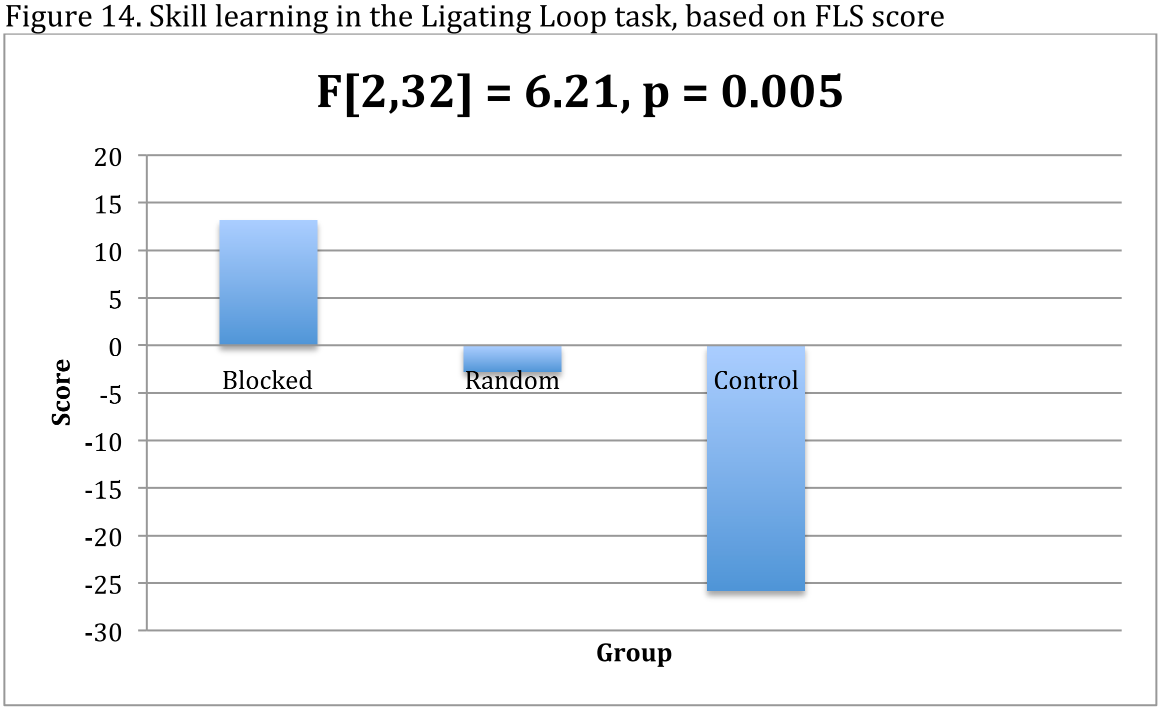
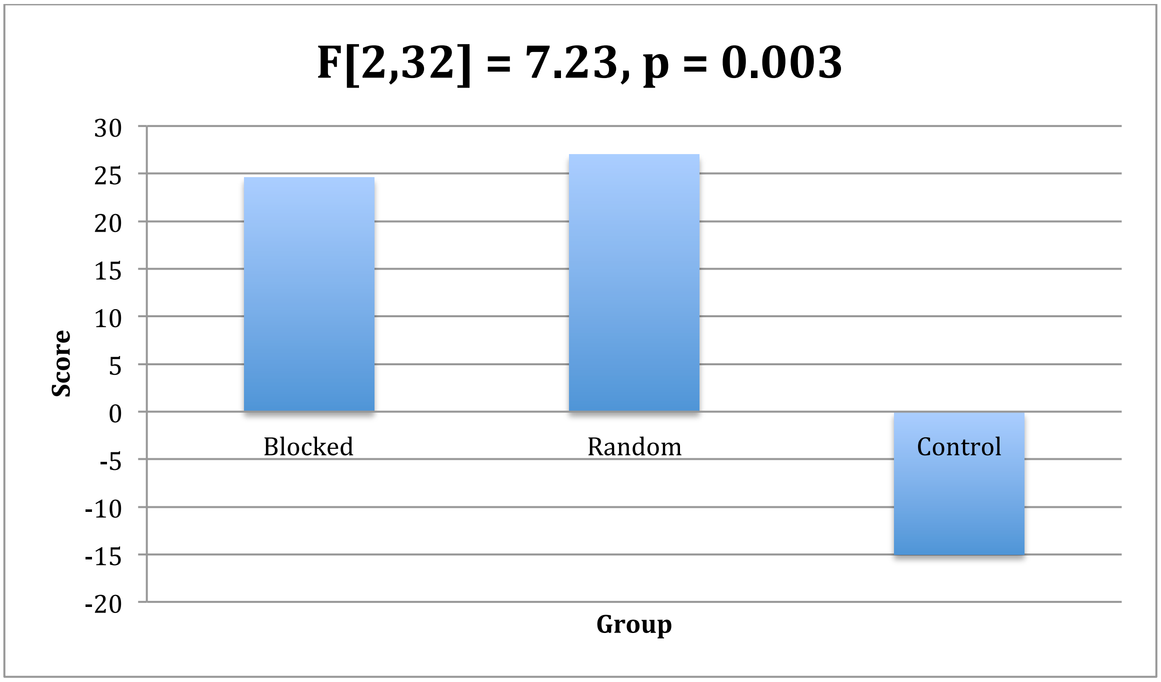


Figure 14. Skill learning in the Ligating Loop task, based on FLS score



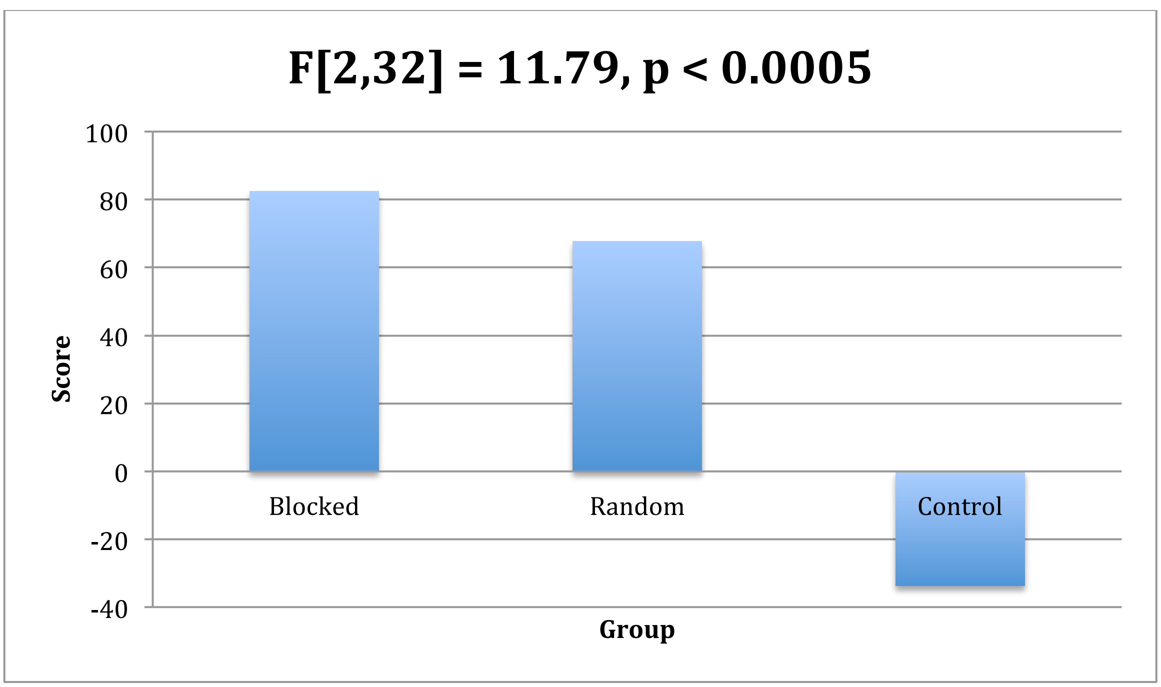
*Post-hoc Tukey's test shows significant difference between blocked and control groups (p = 0.004)

Figure 15. Skill learning in the Intracorporeal Suturing task, based on FLS score



* Post-hoc Tukey's test shows significant difference between the blocked and control groups ($p = 0.006$), and the random and control groups ($p = 0.005$)

Figure 16. Skill learning in the Overall Score, based on FLS score

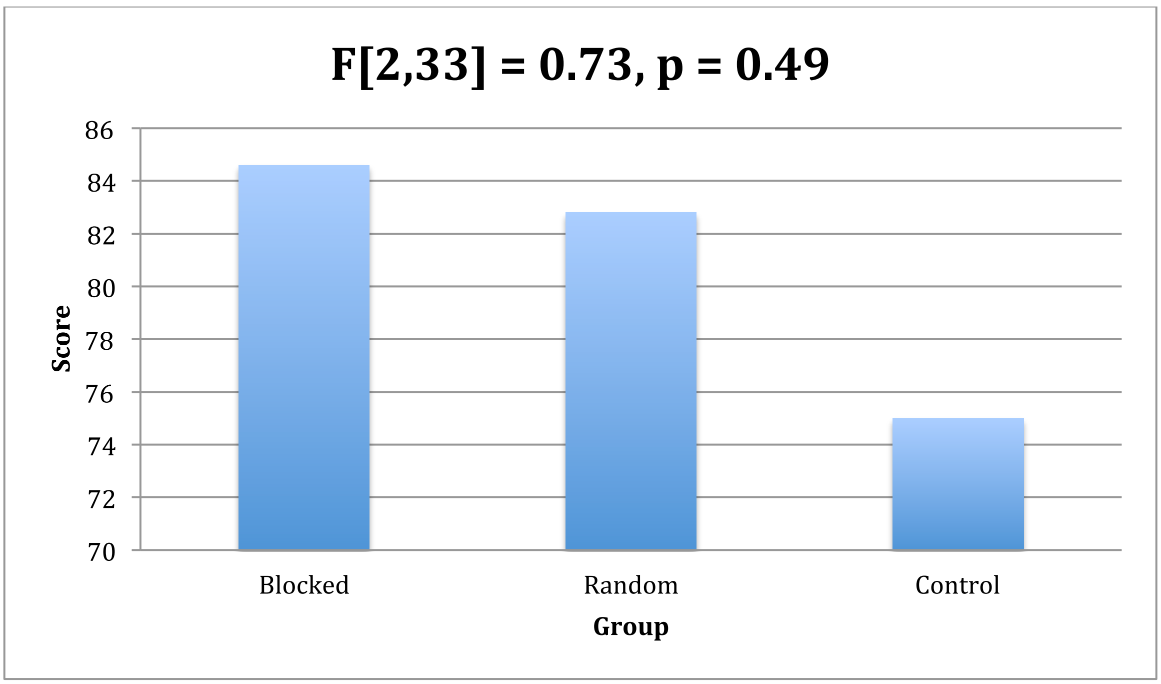


* Post-hoc Tukey's test shows significant difference between the blocked and control groups ($p < 0.0005$), and the random and control groups ($p = 0.001$)

6) Transfer Task

Immediately following the retention tests, all 3 groups performed a transfer of skills task. Figure 17 shows the normalized transfer test scores for all 3 groups. The average of the 2 trials was used to calculate the transfer task score. There appears to be a trend in the training groups attaining a higher score than the control group ($t[34] = 1.20$, $p = 0.24$; independent t test).

Figure 17. Normalized scores in the Transfer of Skills Task



C. Hand Motion Analysis

1) Pretest

Pretest computer-based hand motion efficiency scores were not different between the three groups for number of movements, path-length travelled, and time (see Tables 6-8).

Table 6. Pretest median number of hand movements (1st and 3rd quartiles) for the 3 study groups

	Blocked	Random	Control	Significance Level
Peg Transfer	190.0 (150.4-226.6)	189.8 (159.4-223.3)	180.0 (154.6-225.5)	$\chi^2[2] = 0.31$, p = 0.86
Pattern Cutting	371.5 (322.8-473.5)	381.5 (335.5-421.0)	331.0 (273.3-447.0)	$\chi^2[2] = 0.33$, p = 0.85
Ligating Loop	124.5 (102-136)	113.0 (96.5-136.3)	114.5 (103.8-120.3)	$\chi^2[2] = 0.82$, p = 0.66
IC Suturing	417.0 (331.5-491.8)	455.0 (339.3-539.8)	337.5 (308.0-476.3)	$\chi^2[2] = 1.30$, p = 0.52

* IC = Intracorporeal

Table 7. Pretest median path-length (meters) travelled (1st and 3rd quartiles) for the 3 study groups

	Blocked	Random	Control	Significance Level
Peg Transfer	8.43 (7.45-11.74)	7.91 (7.21-10.13)	8.29 (7.23-8.69)	$\chi^2[2] = 0.38$, p = 0.83
Pattern Cutting	16.42 (9.97-16.95)	13.21 (11.16-14.65)	12.57 (9.0-17.04)	$\chi^2[2] = 0.54$, p = 0.76
Ligating Loop	12.71 (11.17-13.74)	10.95 (10.43-12.51)	11.87 (10.48-13.30)	$\chi^2[2] = 2.62$, p = 0.27
IC Suturing	25.63 (19.72-29.79)	24.99 (19.71-31.33)	20.96 (16.61-27.35)	$\chi^2[2] = 2.03$, p = 0.36

* IC = Intracorporeal

Table 8. Pretest median time in seconds (1st and 3rd quartiles) taken for the 3 study groups

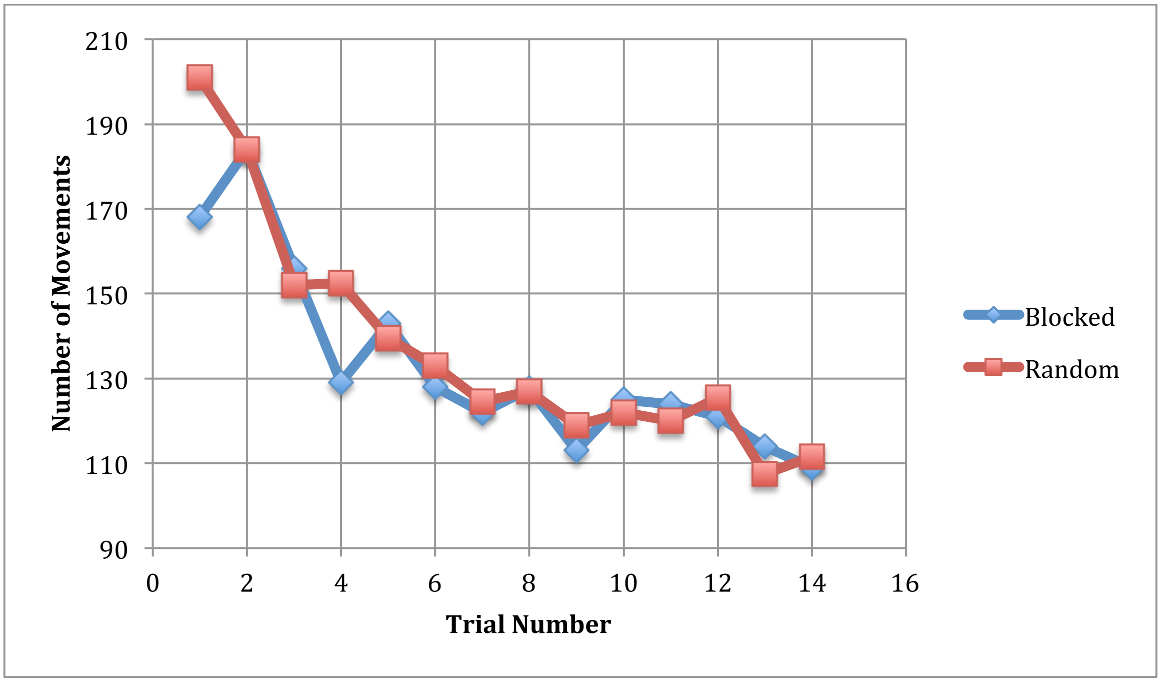
	Blocked	Random	Control	Significance Level
Peg Transfer	141.7 (117.4-186.7)	133.3 (120.1-161.1)	117.9 (109.6-139.5)	$\chi^2[2] = 2.42, p = 0.30$
Pattern Cutting	362.6 (286.3-480.8)	442.2 (362.7-475.3)	433.1 (294.2-489.7)	$\chi^2[2] = 1.15, p = 0.56$
Ligating Loop	116.3 (106.9-133.8)	117.0 (88.0-129.6)	108.0 (93.8-115.9)	$\chi^2[2] = 2.28, p = 0.32$
IC Suturing	383.3 (332.8-495.8)	469.0 (322.5-560.1)	321.3 (301.2-535.5)	$\chi^2[2] = 1.25, p = 0.54$

* IC = Intracorporeal

2) Performance Curves

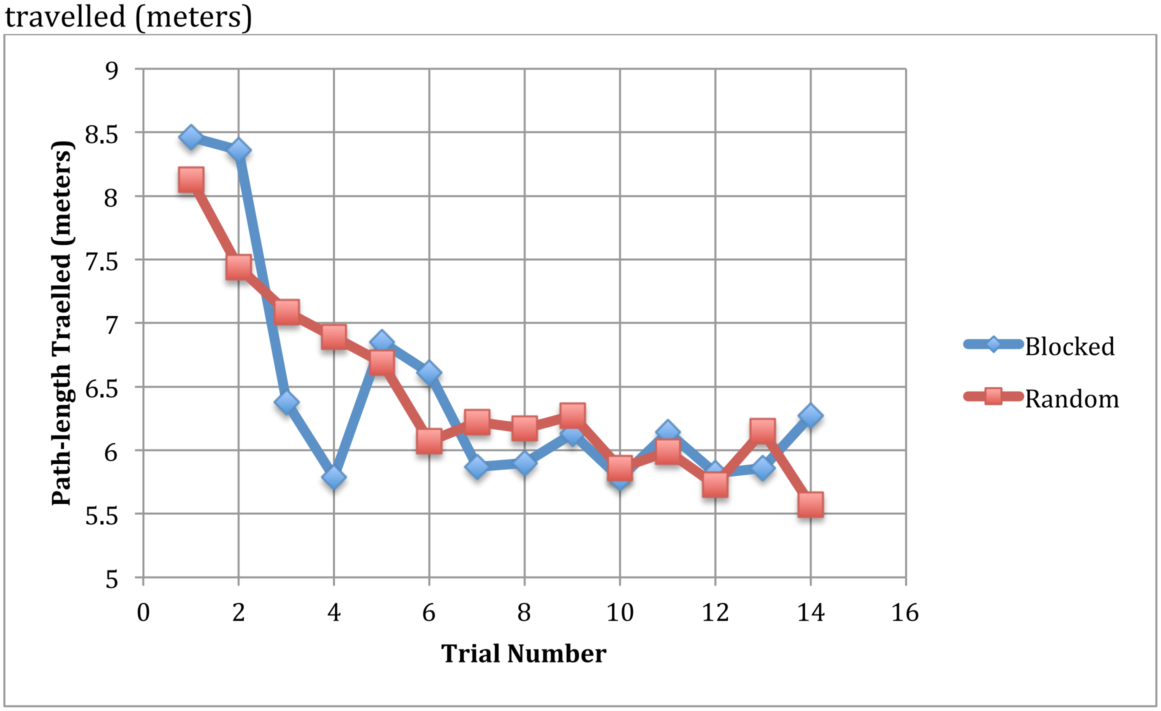
As the participants in the blocked and random groups progressed in their training session, the number of movements, path-length travelled, and time to complete the task improved. The performance curves for all 4 tasks for both groups are shown in Figures 18-29.

Figure 18. Performance curve for the Peg Transfer task median number of movements



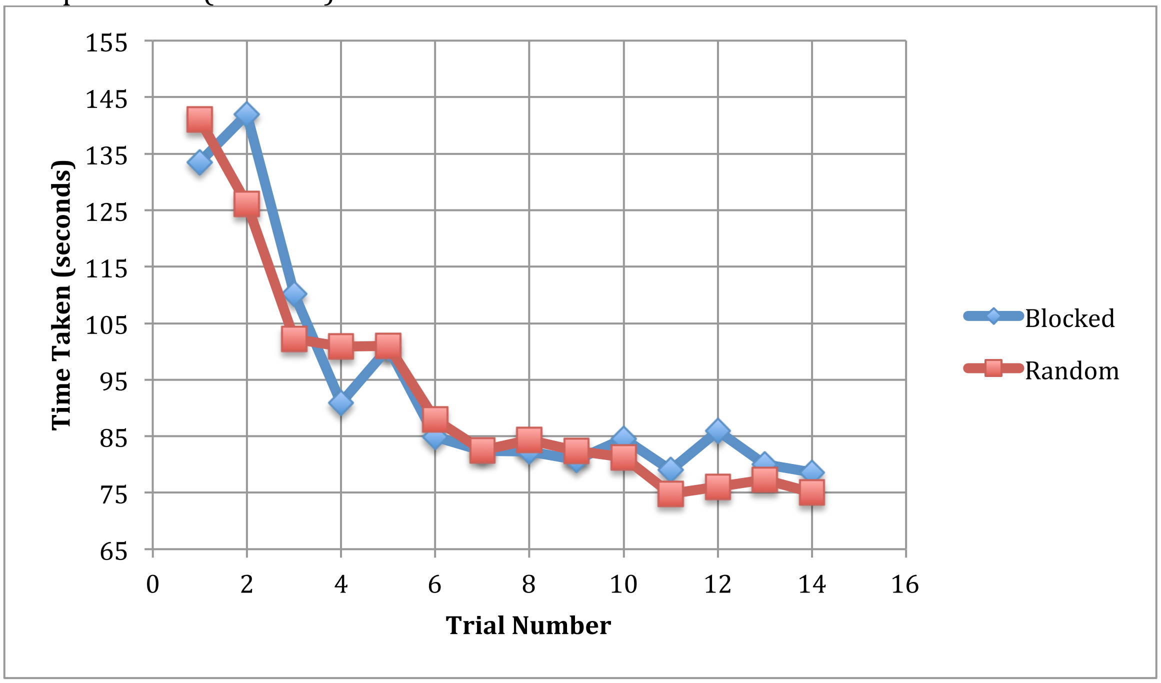
Blocked Post vs. Pretest: Z = 3.06, p=0.002
Random Post vs. Pretest: Z = 3.30, p=0.001

Figure 19. Performance curve for the Peg Transfer task median path-length travelled (meters)



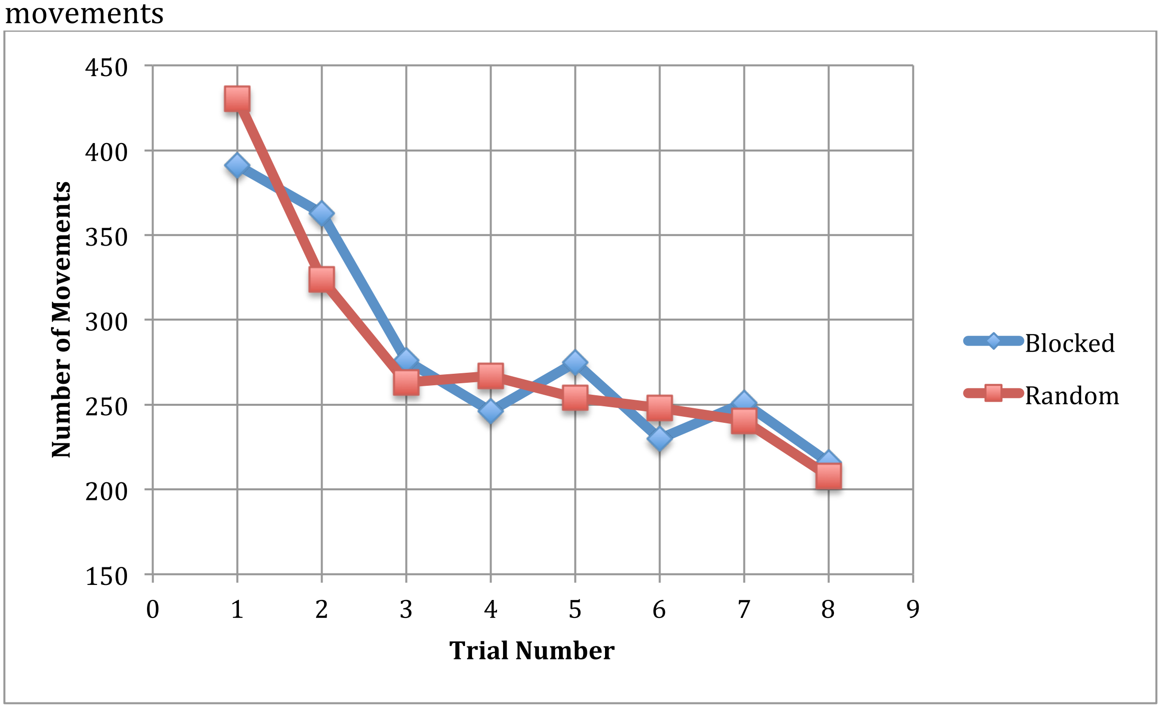
Blocked Post vs. Pretest: Z = 2.82, p=0.005
Random Post vs. Pretest: Z = 3.30, p=0.005

Figure 20. Performance curve for the Peg Transfer task median time taken to complete task (seconds)



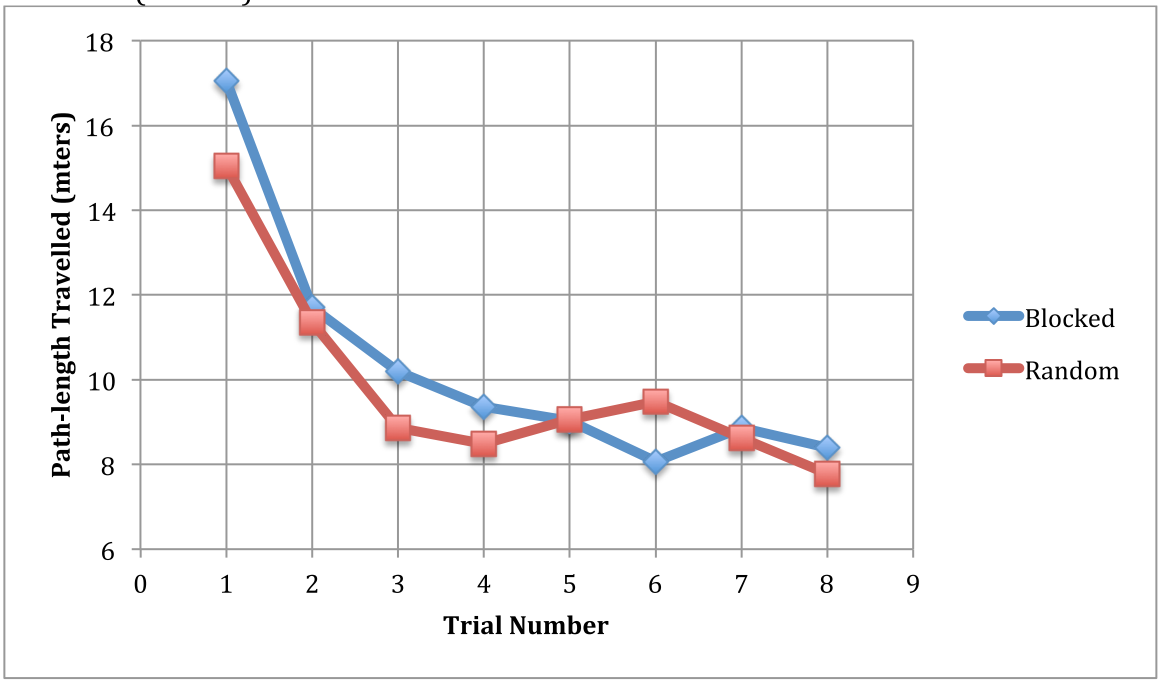
Blocked Post vs. Pretest: Z = 3.06, p=0.002
Random Post vs. Pretest: Z = 3.30, p=0.001

Figure 21. Performance curve for the Pattern Cutting task median number of movements



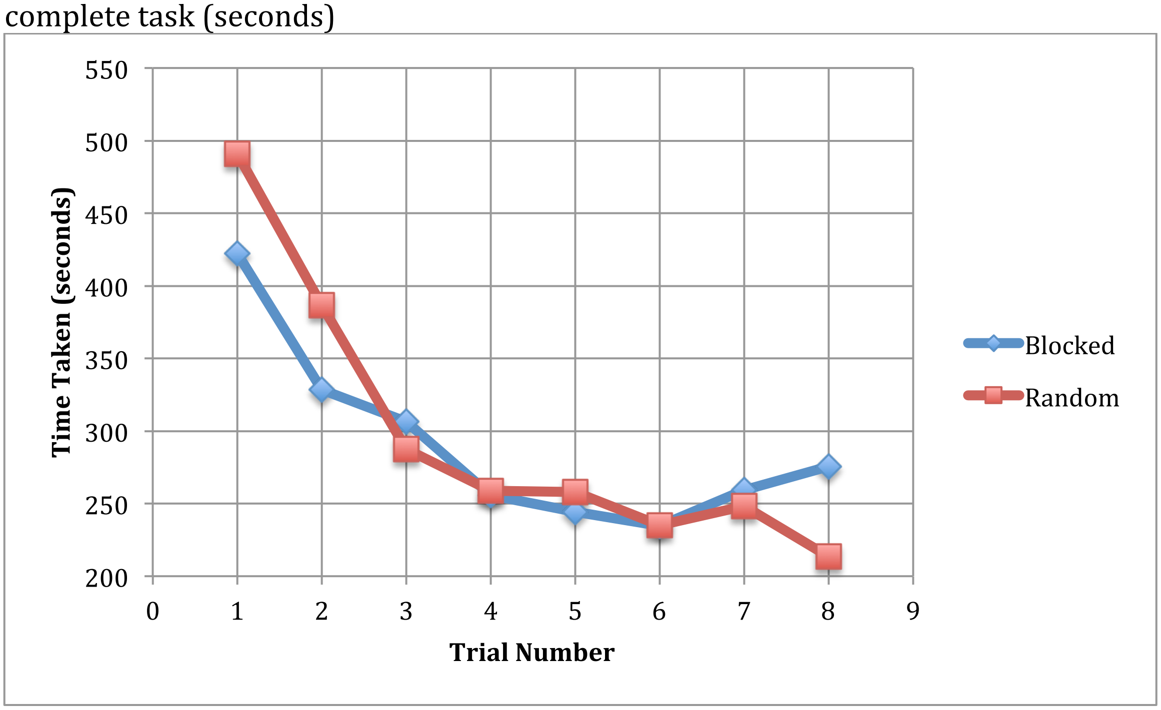
Blocked Post vs. Pretest: Z = 2.83, p=0.005
Random Post vs. Pretest: Z = 3.30, p=0.001

Figure 22. Performance curve for the Pattern Cutting task median path-length travelled (meters)



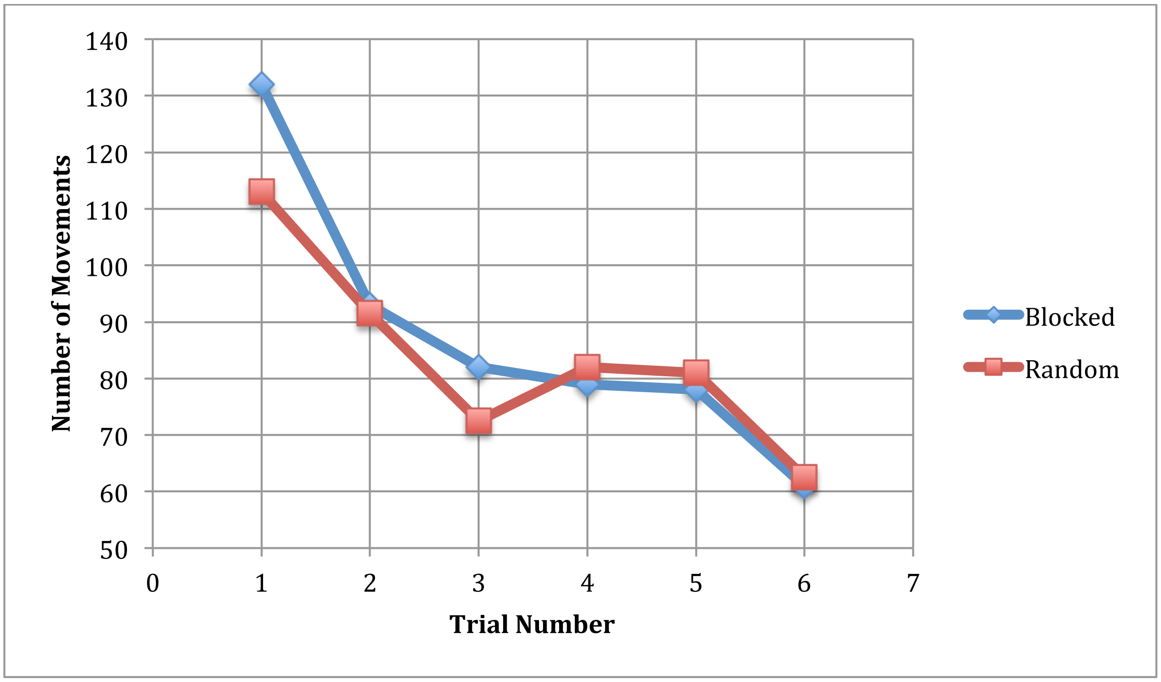
Blocked Post vs. Pretest: Z = 2.83, p=0.005
Random Post vs. Pretest: Z = 3.30, p=0.001

Figure 23. Performance curve for the Pattern Cutting task median time taken to complete task (seconds)



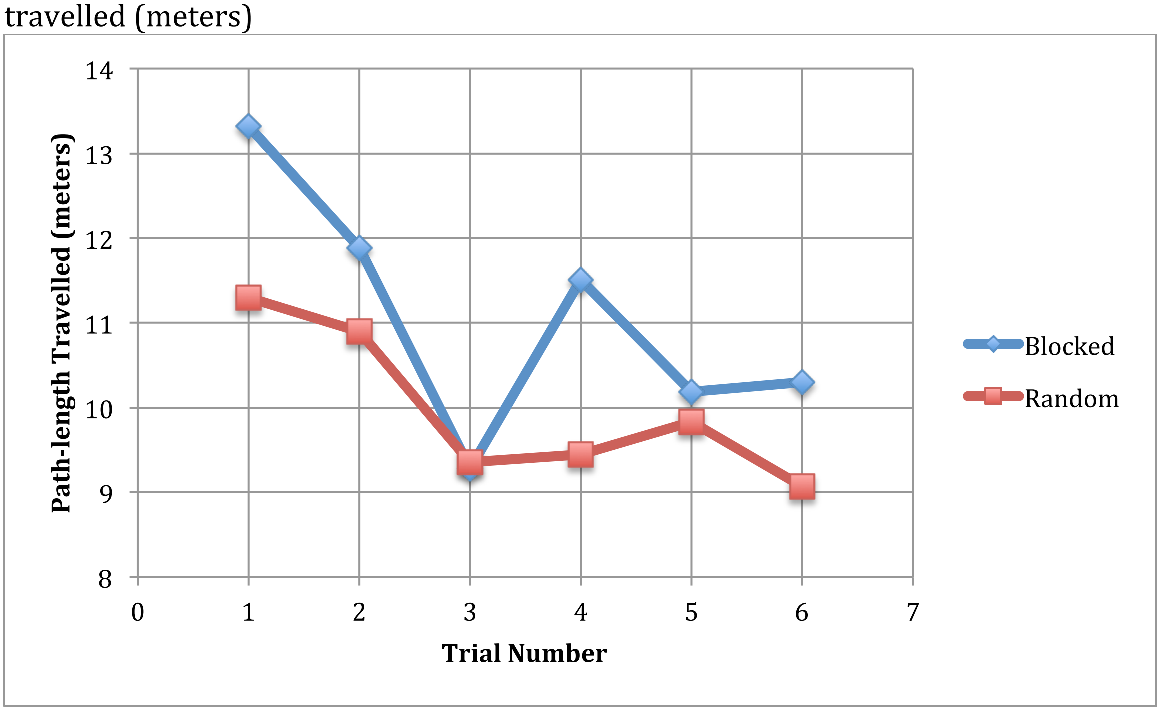
Blocked Post vs. Pretest: Z = 1.92, p=0.055
Random Post vs. Pretest: Z = 3.30, p=0.001

Figure 24. Performance curve for the Ligating Loop task median number of movements



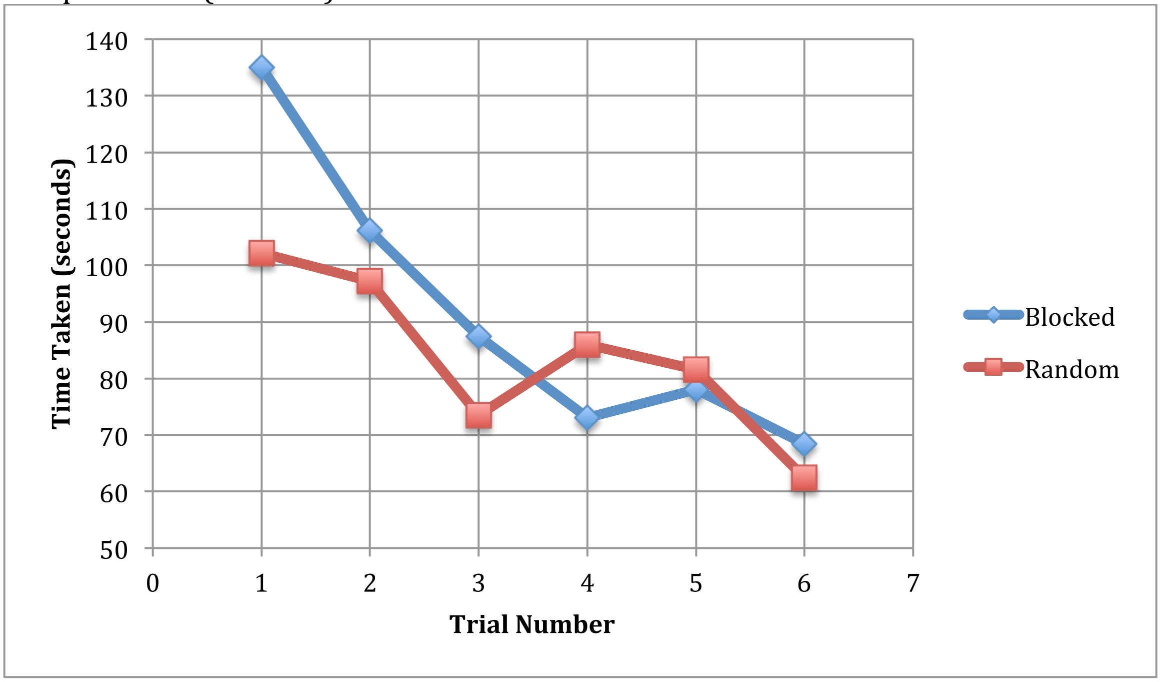
Blocked Post vs. Pretest: Z = 2.90, p=0.004
Random Post vs. Pretest: Z = 3.05, p=0.002

Figure 25. Performance curve for the Ligating Loop task median path-length travelled (meters)



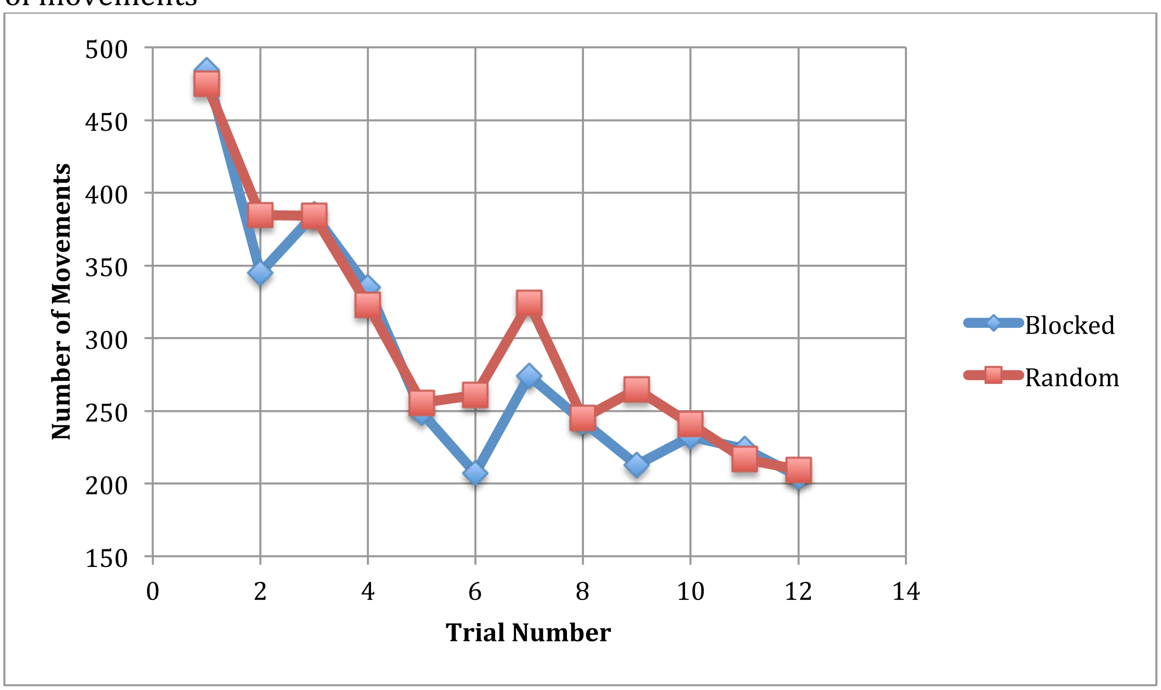
Blocked Post vs. Pretest: Z = 2.69, p=0.007
Random Post vs. Pretest: Z = 2.96, p=0.004

Figure 26. Performance curve for the Ligating Loop task median time taken to complete task (seconds)



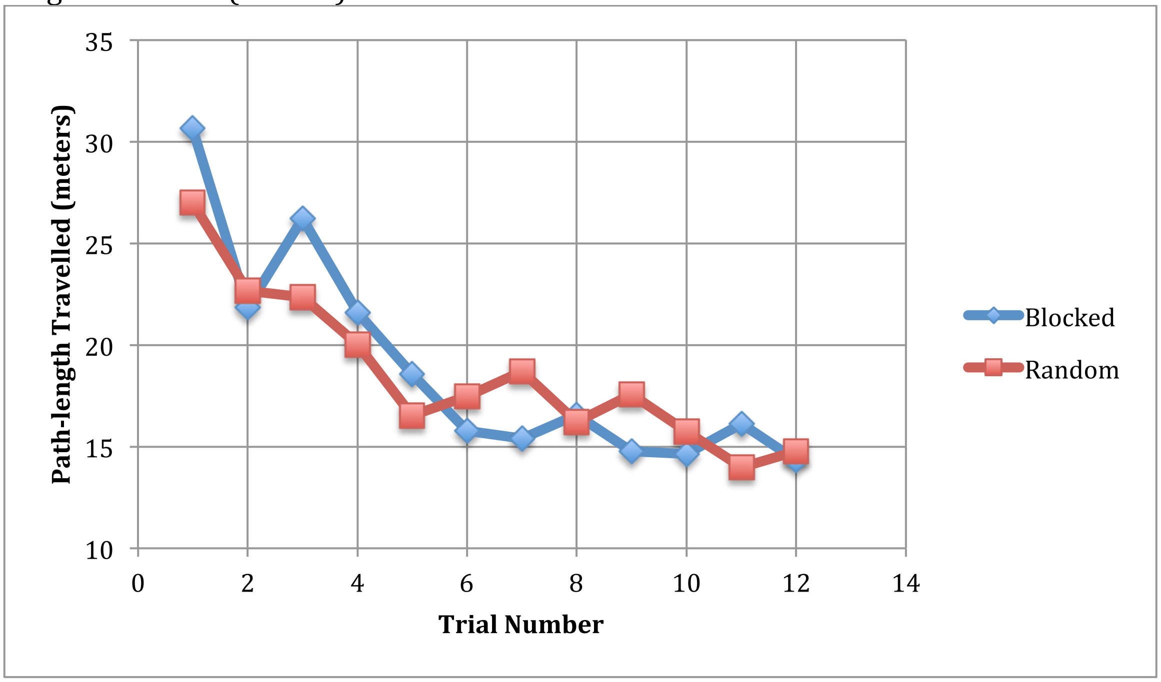
Blocked Post vs. Pretest: Z = 3.18, p=0.001
Random Post vs. Pretest: Z = 3.11, p=0.002

Figure 27. Performance curve for the Intracorporeal Suturing task median number of movements



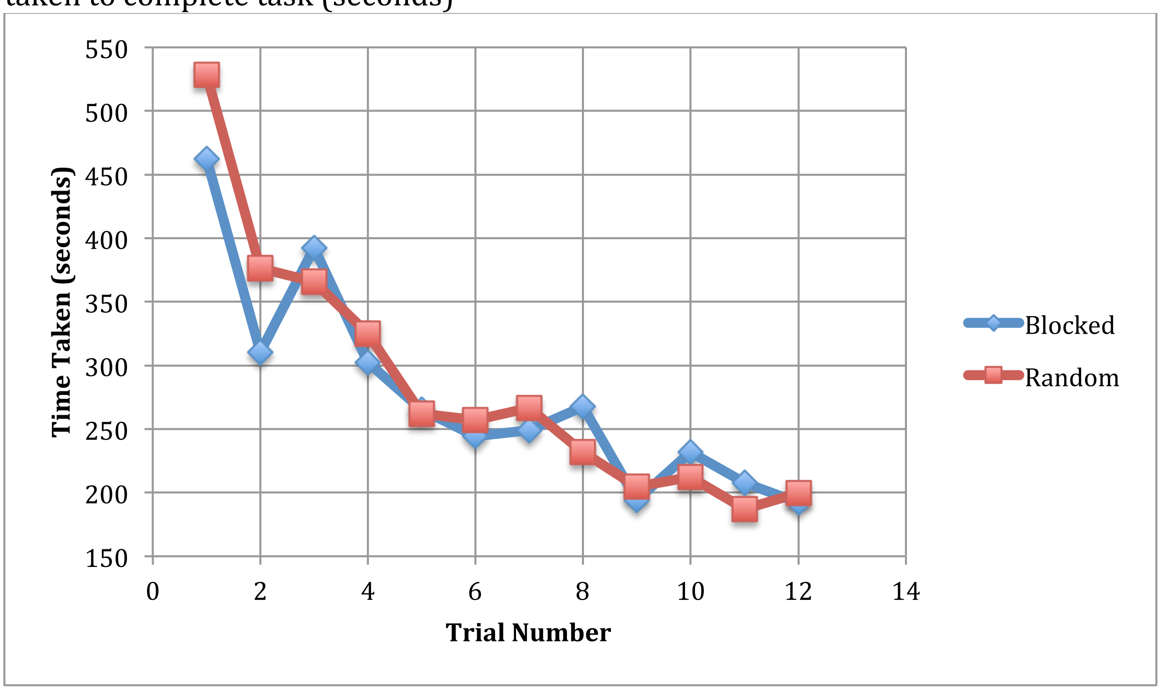
Blocked Post vs. Pretest: Z = 3.18, p=0.001
Random Post vs. Pretest: Z = 2.98, p=0.003

Figure 28. Performance curve for the Intracorporeal Suturing task median path-length travelled (meters)



Blocked Post vs. Pretest: Z = 3.18, p=0.001
Random Post vs. Pretest: Z = 2.67, p=0.008

Figure 29. Performance curve for the Intracorporeal Suturing task median time taken to complete task (seconds)



Blocked Post vs. Pretest: Z = 3.18, p=0.001
Random Post vs. Pretest: Z = 3.05, p=0.002

3) Posttest

The immediate posttest computer-based hand-motion efficiency scores were the same for the number of hand movements and path-length travelled. The only significant difference for the time required to complete the task was between the blocked group and the random group for the Pattern Cutting task, with the random group completing the task faster (see Tables 9-11).

Table 9. Posttest median number of hand movements (1st and 3rd quartiles) for the blocked and random groups

	Blocked	Random	Significance Level
Peg Transfer	114.0 (102.8-127.0)	111.5 (96.0-123.0)	U = 79.5, p = 0.58
Pattern Cutting	216.0 (183.5-291.0)	208.0 (188.5-242.0)	U = 83.5, p = 0.72
Ligating Loop	61.0 (55.5-102.5)	62.5 (54.5-83.8)	U = 81.0, p = 0.63
IC Suturing	205.0 (175.5-253.5)	209.5 (178.8-330.5)	U = 77.5, p = 0.51

* IC = Intracorporeal

Table 10. Posttest median path-length (meters) travelled (1st and 3rd quartiles) for the blocked and random groups

	Blocked	Random	Significance Level
Peg Transfer	6.30 (4.80-7.85)	5.57 (4.76-6.89)	U = 76.0, p = 0.47
Pattern Cutting	8.39 (6.06-10.75)	7.78 (5.86-11.10)	U = 81.0, p = 0.63
Ligating Loop	10.30 (8.46-11.64)	9.07 (8.52-10.14)	U = 59.0, p = 0.12
IC Suturing	14.34 (12.29-15.60)	14.78 (12.30-20.39)	U = 74.0, p = 0.41

* IC = Intracorporeal

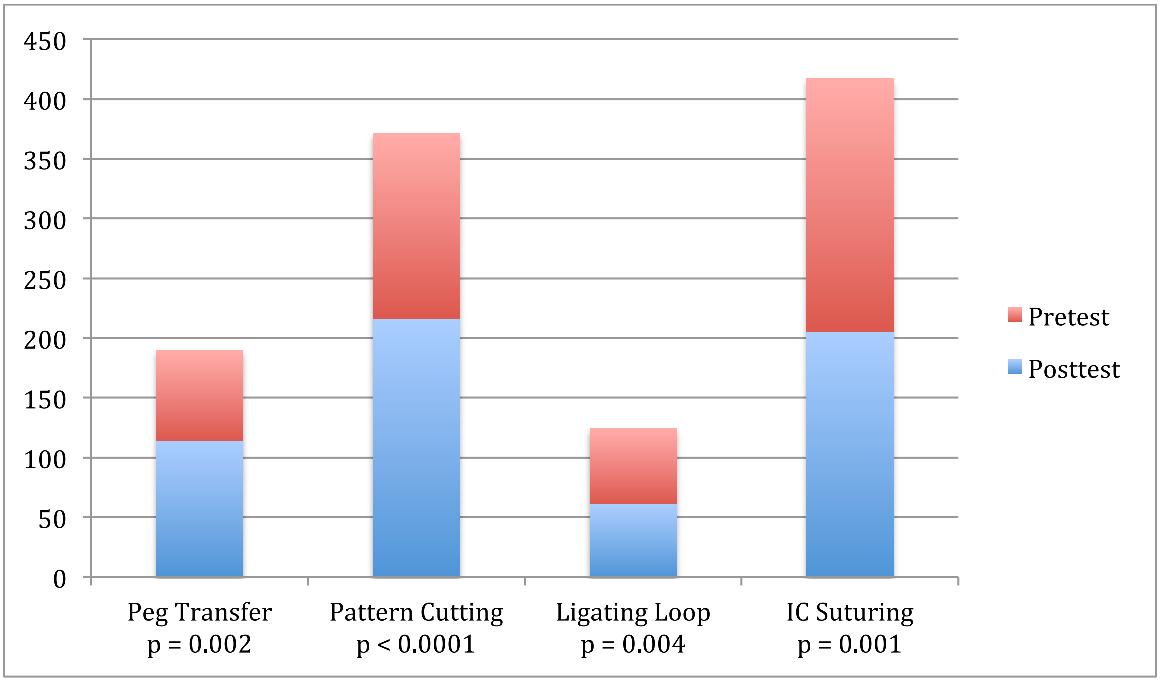
Table 11. Posttest median time in seconds (1st and 3rd quartiles) taken for the blocked and random groups

	Blocked	Random	Significance Level
Peg Transfer	79.6 (74.0-86.7)	75.0 (65.8-75.0)	U = 68.6, p = 0.28
Pattern Cutting	275.5 (201.8-293.3)	213.4 (176.7-245.9)	U = 50.0, p = 0.047
Ligating Loop	68.4 (59.1-99.5)	62.5 (58.3-74.5)	U = 72.0, p = 0.36
IC Suturing	192.1 (183.4-216.8)	199.9 (174.5-305.5)	U = 77.0, p = 0.50

*IC = Intracorporeal

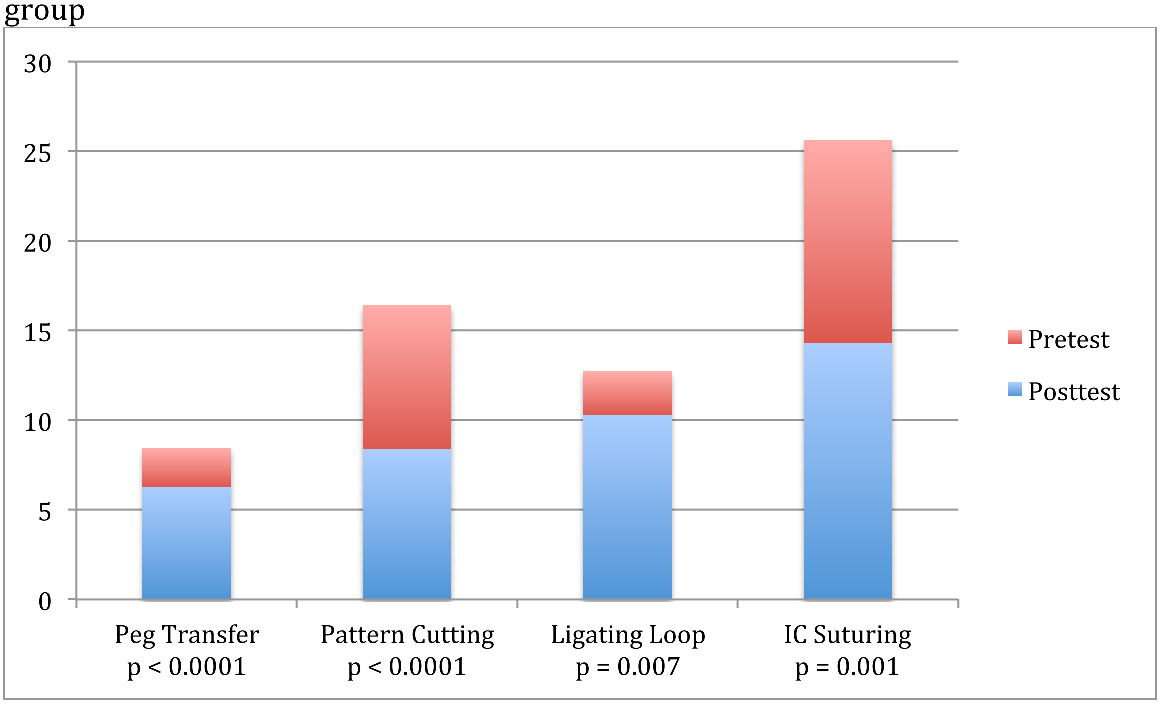
The improvement in the computer-based hand-motion analysis scores within the blocked and random groups from pretest to posttest was assessed using the Wilcoxon signed ranks test. The results are shown in Figures 30-35. Both groups improved in all measures except for the time taken to complete the pattern cutting task, where the blocked group showed no improvement.

Figure 30. Median number of movements for the 4 tasks in the Blocked group



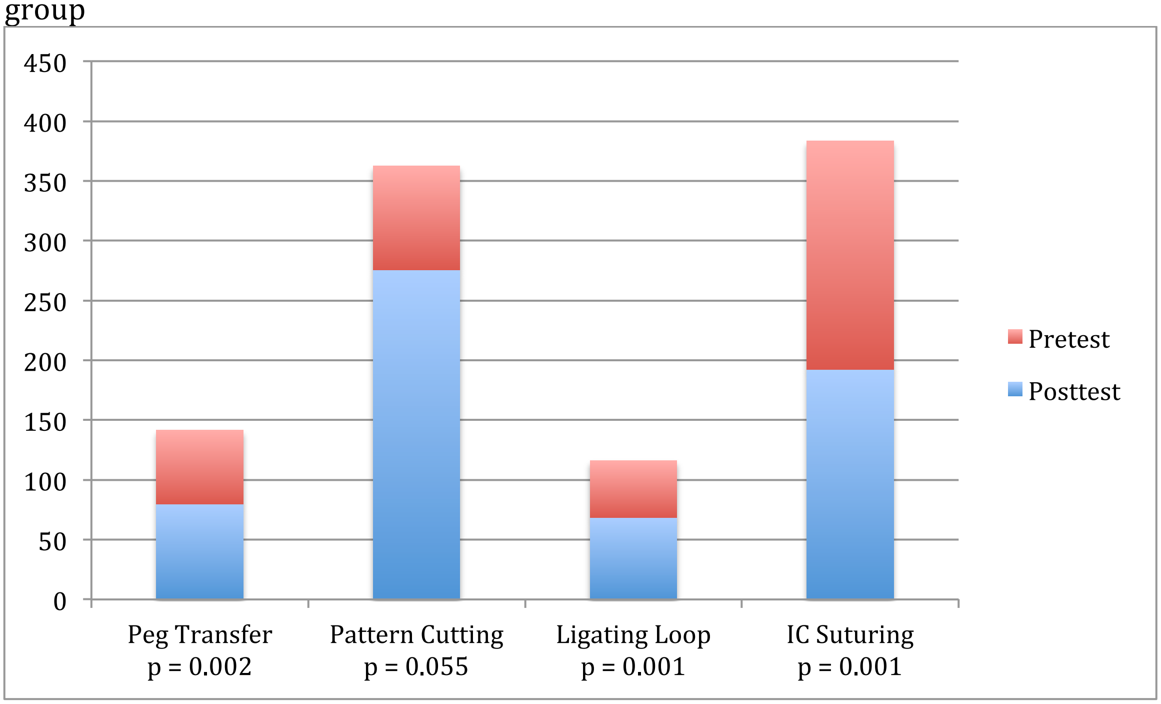
* IC = Intracorporeal

Figure 31. Median path-length travelled (meters) for the 4 tasks in the Blocked group



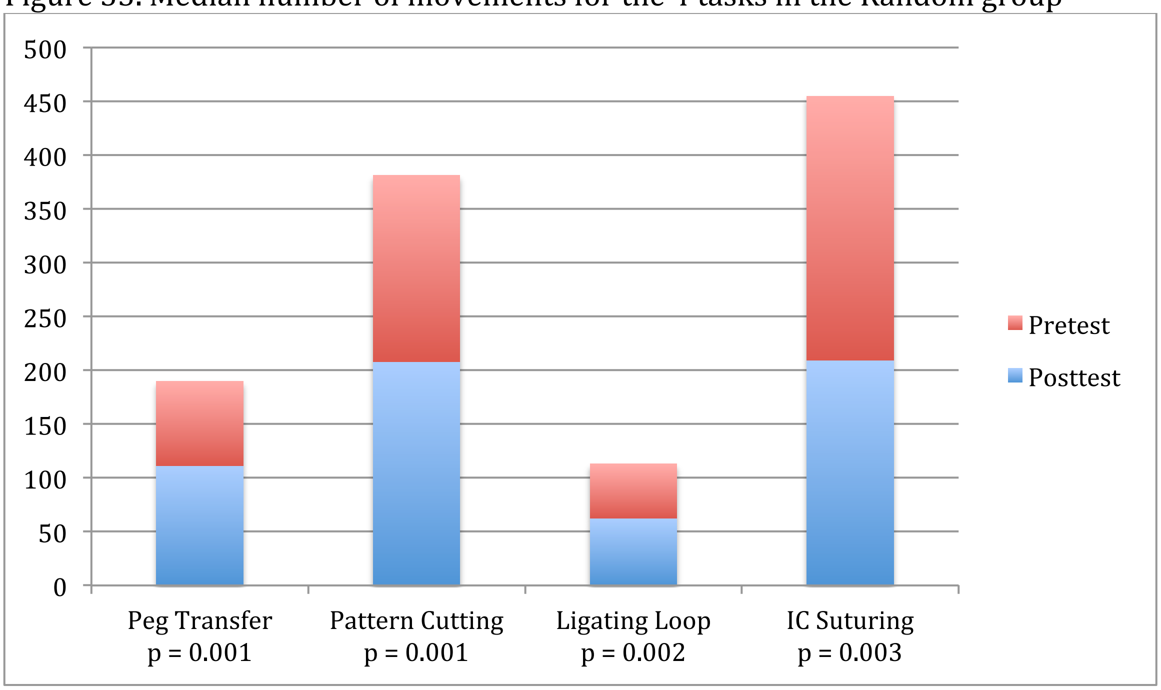
* IC = Intracorporeal

Figure 32. Median time (seconds) taken to complete the 4 tasks in the Blocked group



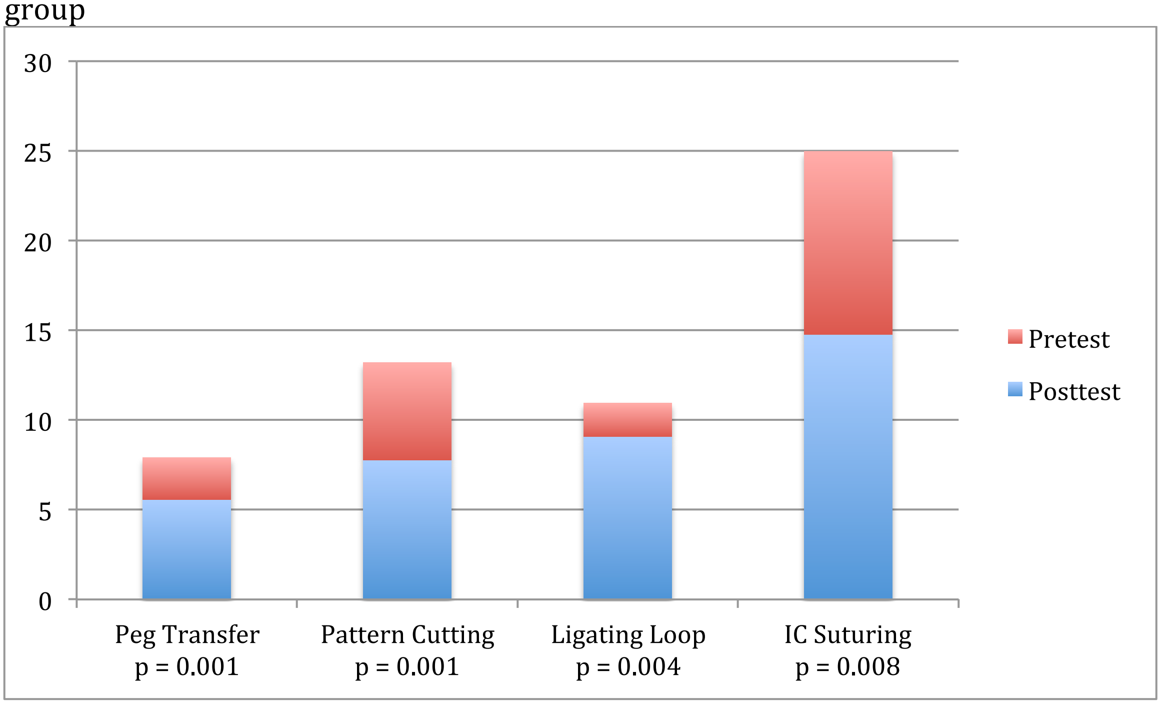
* IC = Intracorporeal

Figure 33. Median number of movements for the 4 tasks in the Random group



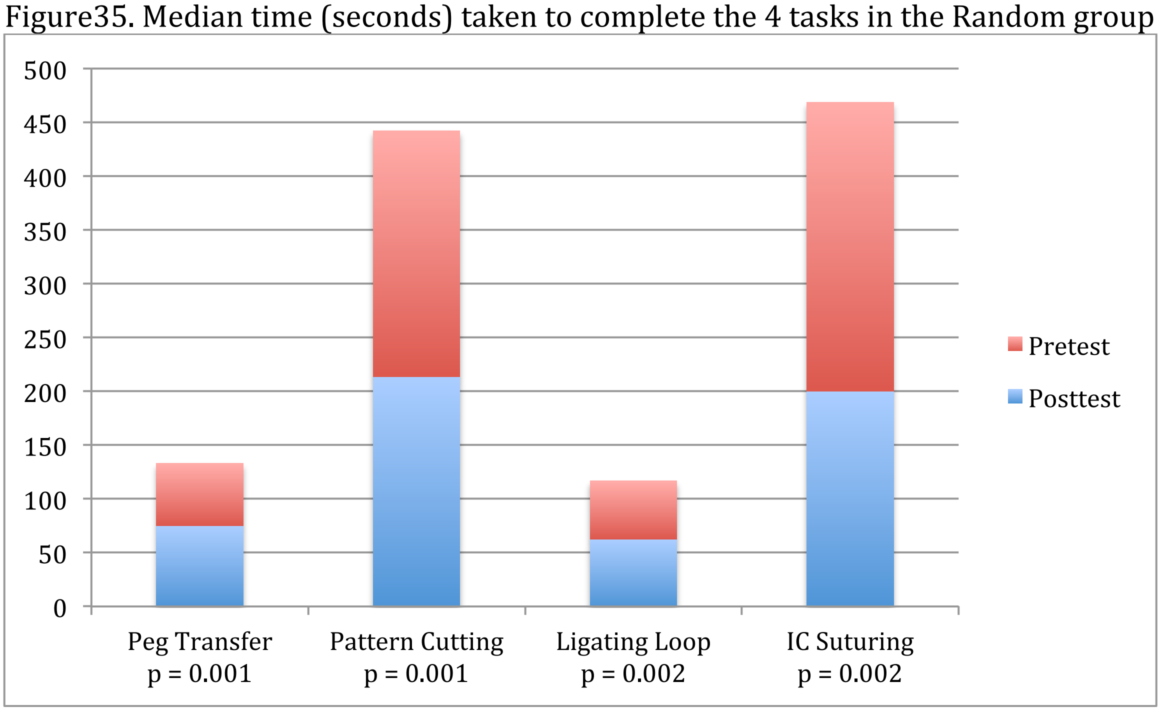
* IC = Intracorporeal

Figure 34. Median path-length travelled (meters) for the 4 tasks in the Random group



* IC = Intracorporeal

Figure35. Median time (seconds) taken to complete the 4 tasks in the Random group



* IC = Intracorporeal

4) Skill Retention

After completing the pretest and training sessions, all 3 groups returned 5-6 weeks later to perform retention tests on all 4 tasks. Tables 12-14 show the Wilcoxon signed ranks test results for the number of movements, path-length travelled, and time taken to complete the tasks for the blocked group. It can be seen that the number of hand movements shows a significant increase only for the suturing task. The path-length travelled increased significantly for the Ligating Loop and Suturing task. Finally, for the time taken to complete the task, only the Pattern Cutting group did not show a significant increase. It can be seen that the IC suturing task showed skill degradation in all three efficiency scores for the blocked group.

Table 12. Comparison of retention and posttest median number of movements (1st and 3rd quartiles) for the Blocked group

	Posttest	Retention Test	Significance Level
Peg Transfer	114 (102.8-127.0)	138 (119.5-152.0)	z = 1.57, p = 0.12
Pattern Cutting	216 (183.5-291.0)	243 (185-261.5)	z = 0.94, p = 0.35
Ligating Loop	61 (55.5-102.5)	92 (74.0-117.0)	z = 1.49, p = 0.14
IC Suturing	205 (175.5-253.5)	277 (198.5-352.5)	z = 2.35, p = 0.019

* IC = Intracorporeal

Table 13. Comparison of retention and posttest median path-length (meters) travelled (1st and 3rd quartiles) for the Blocked group

	Posttest	Retention Test	Significance Level
Peg Transfer	6.30 (4.80-7.85)	6.83 (5.81-7.64)	z = 1.49, p = 0.14
Pattern Cutting	8.39 (6.06-10.75)	7.74 (6.78-9.61)	z = 0.80, p = 0.42
Ligating Loop	10.30 (8.46-11.64)	13.07 (12.12-15.24)	z = 2.48, p = 0.013
IC Suturing	14.34 (12.29-15.60)	19.49 (16.49-22.79)	z = 2.69, p = 0.007

* IC = Intracorporeal

Table 14. Comparison of retention and posttest median time in seconds to complete the task (1st and 3rd quartiles) for the Blocked group

	Posttest	Retention Test	Significance Level
Peg Transfer	79.6 (74.0-86.7)	98 (93.0-108.6)	z = 2.20, p = 0.028
Pattern Cutting	275.5 (201.8-293.3)	259.8 (196.5-323.1)	z = 0.04, p = 0.97
Ligating Loop	68.4 (59.1-99.5)	95.2 (78.9-121.2)	z = 2.20, p = 0.028
IC Suturing	192.1 (183.4-216.8)	254.3 (201.8-345.6)	z = 2.62, p = 0.009

* IC = Intracorporeal

Tables 15-17 show the Wilcoxon signed ranks test results for the number of movements, path-length travelled, and time taken to complete the tasks for the random group. The Peg Transfer task and the Ligating Loop task showed a significant increase in all 3 hand-motion efficiency scores. Additionally, the Pattern Cutting task showed an increase in time taken to complete the task. The Random group was robust to skill loss for all the efficiency scores for the suturing task.

Table 15. Comparison of retention and posttest median number of movements (1st and 3rd quartiles) for the Random group

	Posttest	Retention Test	Significance Level
Peg Transfer	111.5 (96.0-123.0)	126.6 (106.8-151.8)	z = 2.54, p = 0.011
Pattern Cutting	208 (188.5-242.0)	266.5 (222.0-372.5)	z = 1.85, p = 0.06
Ligating Loop	62.5 (54.5-83.8)	128.5 (72-172.5)	z = 2.86, p = 0.004
IC Suturing	209.5 (178.8-330.5)	286.5 (205.3-341.8)	z = 1.22, p = 0.22

* IC = Intracorporeal

Table 16. Comparison of retention and posttest median path-length (meters) travelled (1st and 3rd quartiles) for the Random group

	Posttest	Retention Test	Significance Level
Peg Transfer	5.57 (4.76-6.89)	6.62 (5.83-8.02)	z = 2.67, p = 0.008
Pattern Cutting	7.78 (5.86-11.10)	10.21 (7.04-13.11)	z = 1.85, p = 0.06
Ligating Loop	9.07 (8.52-10.14)	12.60 (11.21-16.78)	z = 3.17, p = 0.002
IC Suturing	14.78 (12.30-20.39)	20.75 (13.78-23.27)	z = 1.41, p = 0.16

* IC = Intracorporeal

Table 17. Comparison of retention and posttest median time in seconds to complete the task (1st and 3rd quartiles) for the Random group

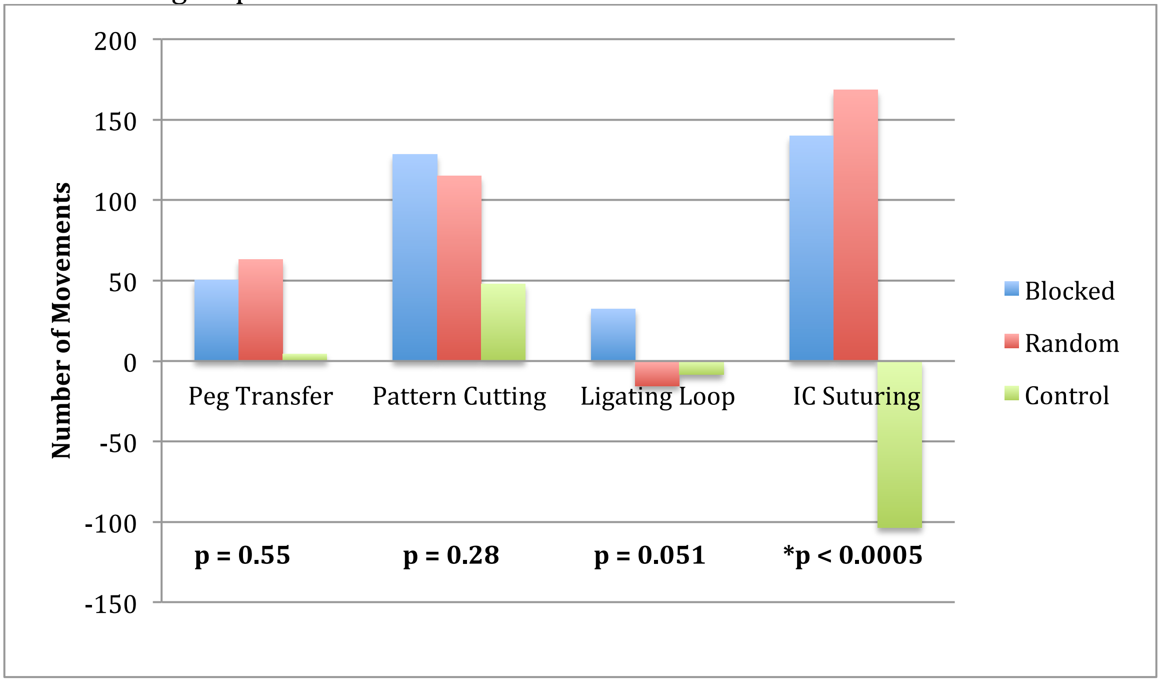
	Posttest	Retention Test	Significance Level
Peg Transfer	75 (65.8-81.0)	85.8 (71.7-114.0)	z = 2.70, p = 0.007
Pattern Cutting	213.4 (176.7-245.9)	273.4 (231.6-358.0)	z = 2.54, p = 0.011
Ligating Loop	62.5 (58.3-74.5)	116.1 (86.7-142.5)	z = 3.05, p = 0.002
IC Suturing	199.9 (174.5-305.5)	252.0 (181.2-316.9)	z = 1.10, p = 0.27

* IC = Intracorporeal

5) Skill Learning

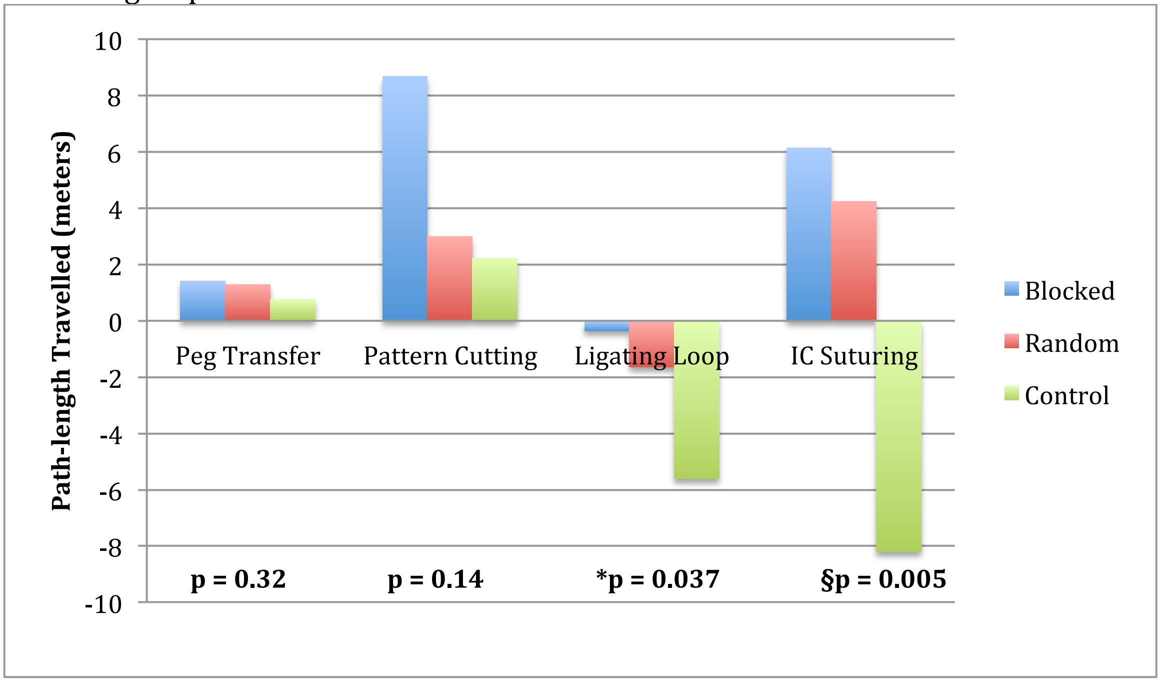
The amount of skill learning (assessed as the improvement in motion efficiency scores between the retention test and the pretest scores) for each of the 3 hand-motion efficiency scores for all 4 tasks is shown in Figures 36-38. For the ligating loop and intracorporeal suturing tasks differences were seen between the practice groups (blocked and random) and the control group, but not between the blocked and random groups. No differences were observed between the groups for the peg transfer and pattern cutting tasks.

Figure 36. Change in median number of movements from pretest to retention test for all three groups in all four tasks



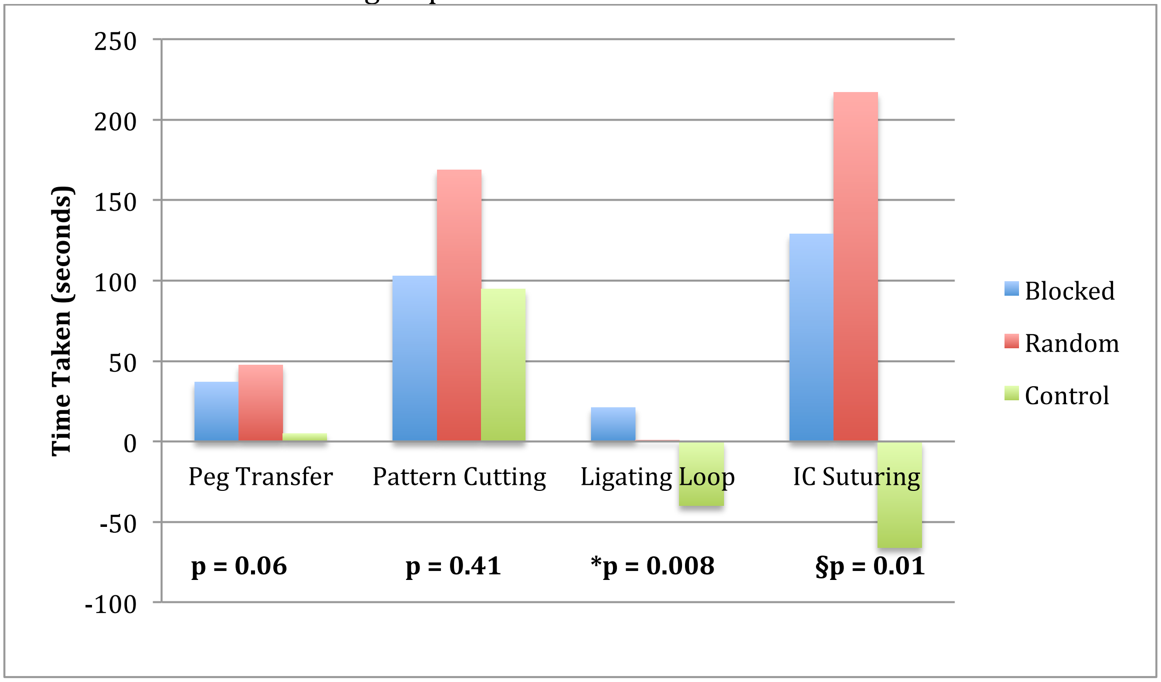
* Post-hoc Mann-Whitney U test for IC Suturing shows significant difference between blocked and control groups ($p < 0.0005$) as well as the random and control groups ($p < 0.0005$)

Figure 37. Change in median path-length travelled from pretest to retention test for all three groups in all four tasks



*Post-hoc Mann-Whitney U test for Ligating Loop shows significant difference between blocked and control group (p = 0.004)
§Post-hoc Mann-Whitney U test for Suturing shows significant difference between blocked and control group (p = 0.009) as well as the random and control groups (p = 0.001)

Figure 38. Change in median time taken to complete the task from pretest to retention test for all three groups in all four tasks



* Post-hoc Mann-Whitney U test for Ligating Loop shows significant difference between blocked and control group (p = 0.002) as well as the random and control groups (p = 0.044)
§ Post-hoc Mann-Whitney U test for Suturing shows significant difference between blocked and control group (p = 0.03) as well as the random and control groups (p = 0.003)

6) Transfer Task

Immediately following the retention tests, all 3 groups performed a transfer of skills task. Figures 39-41 show the hand-motion efficiency scores for the number of movements, path-length travelled, and time taken to complete the task. No significant differences were observed by the Kruskal-Wallis test. Similar to the FLS scores, there appears to be a trend to better efficiency scores when comparing groups that underwent training (blocked and random) to the control group (Mann-Whitney U test: number of movements, p = 0.27; path-length travelled, p = 0.16; time, p = 0.27).

Figure 39. Median number of movements for all three groups in the Transfer of Skills task

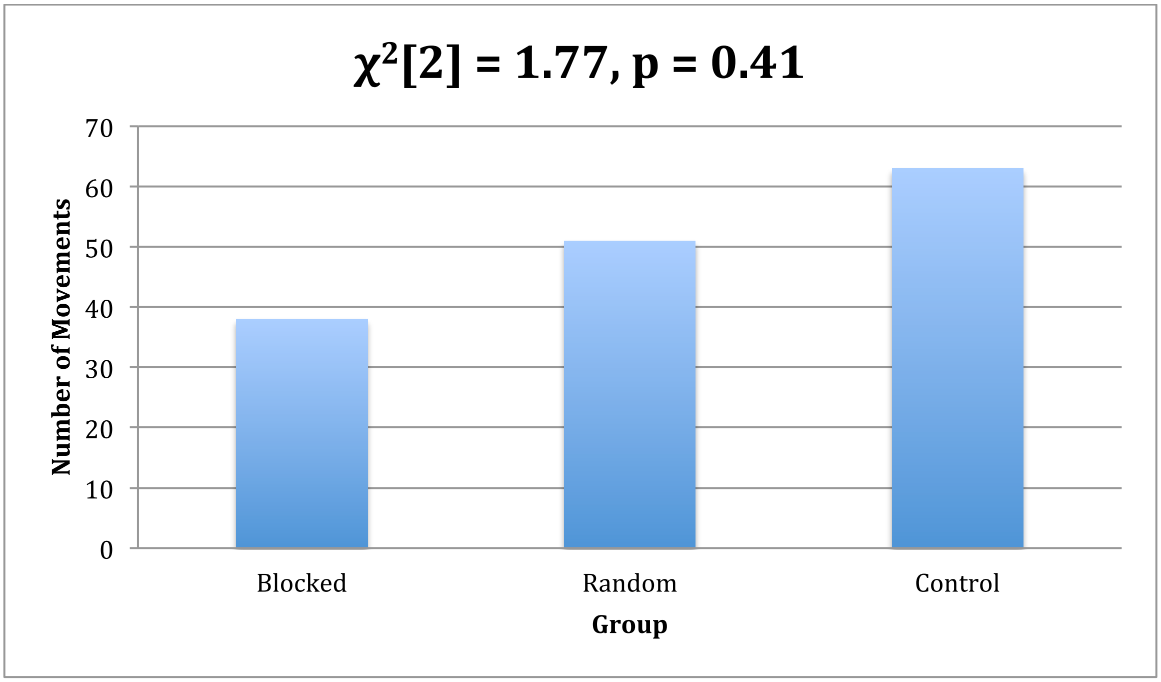


Figure 40. Median path-length travelled (meters) for all three groups in the Transfer of Skills task

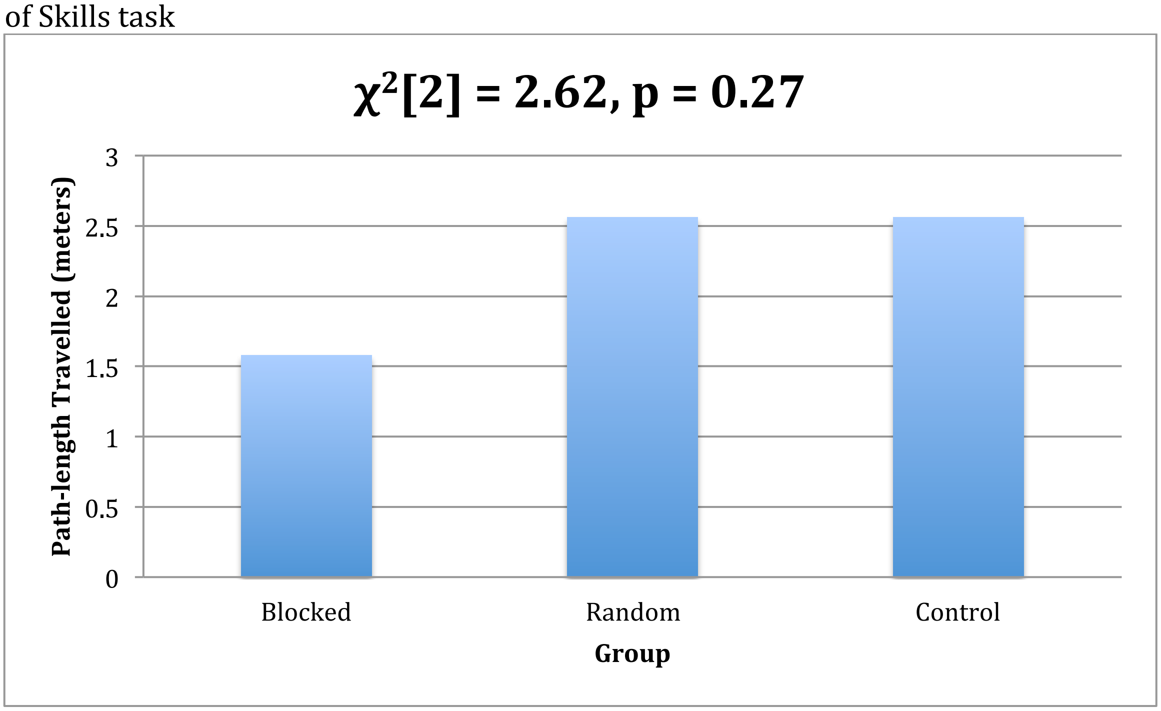
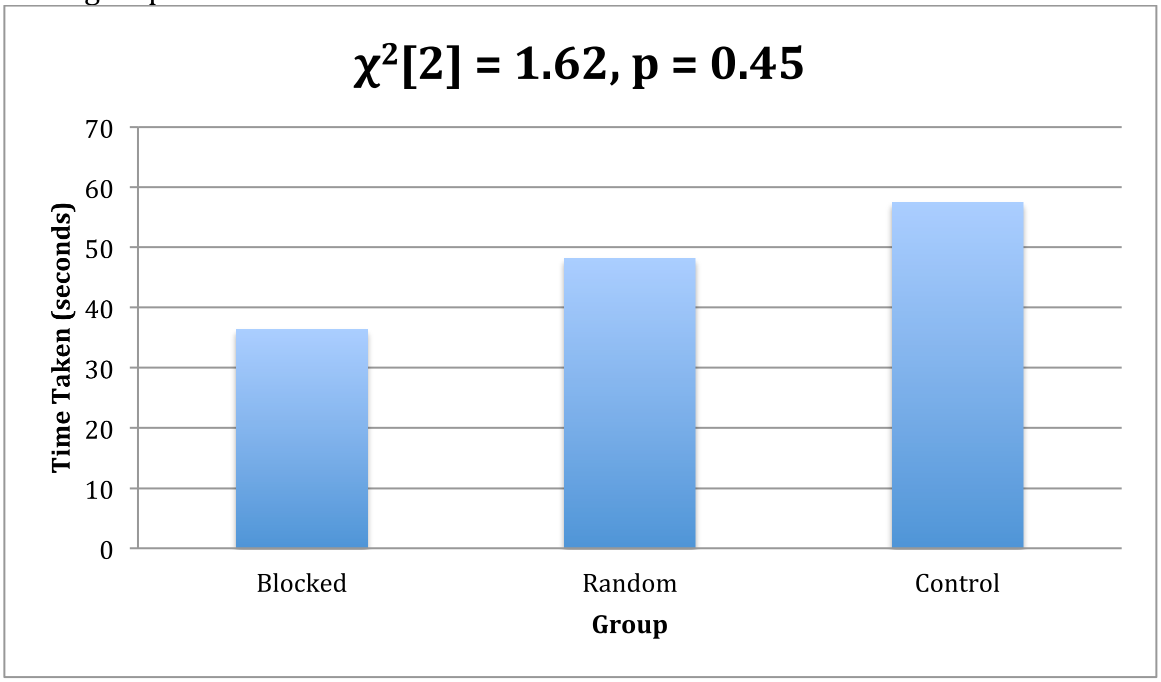


Figure 41. Median time taken (seconds) to complete the Transfer of Skills task for all three groups



Chapter 8. Discussion

Immediately following training, the blocked and random groups completed a posttest for each of the 4 tasks. Paired t-tests showed a positive effect of training for both blocked and random groups for the FLS scores for all tasks. However, there was no statistically significant difference between training groups in posttest scores (see Table 3). There was a trend towards significance for improvement in score for the suturing task with the blocked group recording a more improved score (43.2 vs. 34.4, $p = 0.13$). The hand-motion efficiency scores similarly showed a positive effect of training for both the blocked and random training groups for all four tasks except on the time to complete the pattern-cutting task for the blocked group. Again, there was no statistically significant difference between training groups in posttest scores except for the time taken to complete the pattern-cutting task, which had a statistically significant difference between the blocked and random group, with the random group completing the task faster. These results are in contrast to the motor learning literature where it is suggested that practicing a task in repeated (blocked) fashion allows the trainee to obtain superior performance immediately following training as compared to practicing a task in a random order. Several explanations can account for this result, which will be discussed in the following sections.

When the acquisition curves for the computer-based hand motion efficiency scores are inspected (Figures 18-29) we can see mixed results. However, overall the median number of movements, path-length travelled, and time taken to complete the tasks appeared similar for all four tasks.

In a study comparing FLS scores with those of the computer-based hand motion efficiency scores, Xeroulis *et al.*[38] separated participants into three different skills levels (junior resident, senior resident, and staff surgeon) and found construct validity for the number of movements for the peg transfer, ligating loop, and suturing tasks. The total path-length travelled showed construct validity for the peg transfer task and the suturing task. Total time taken to complete the task showed construct validity for all four tasks. This is not surprising since time is the main component of the FLS scores. They also found a statistically significant correlation between total FLS scores and all three motion efficiency scores. Taken together it appears as though the number of hand movements and total time are the more important motion efficiency metrics. If we apply the motion efficiency metrics that can successfully distinguish skill level back to our results, we see that overall the blocked and random groups perform similarly.

With time comprising the majority of the FLS score and the only motion efficiency score that showed construct validity for all four tasks, the question arises as to why we don't just use time when assessing surgical performance. This question was answered by Datta *et al*. by performing partial correlation tests[37]. When controlling for time, the number of movements made was found to be statistically significant with respect to surgical experience. However, when controlling for movement, the time taken to complete the tasks had no relationship with surgical experience. This suggests that operative speed is secondary to the efficiency of hand movement. It is for this reason that we included several outcome measures with multiple components into our study. In the clinical setting the fastest surgeon is not necessarily the best surgeon.

The main principle behind hand-motion analysis systems, such as the ICSAD being used as part of surgical education is the thought that as motor skill is learned, the movements made by the trainee will eventually transition from being highly inaccurate and uncoordinated to becoming more efficient, more coordinated, smoother, and require fewer movements to perform the same task[77]. Construct validity has been previously shown for the ICSAD for both laparoscopic[34] and open surgical simulation[36, 78], as well as for an entire simulated laparoscopic surgical procedure[79].

Our results seem to be consistent with other studies on assessing practice schedules on surgical tasks. Dubrowski *et al*. found no statistically significant difference in post-test scores between the random and blocked groups for 5 separate bone-plating tasks[17]. Similarly, Brydges *et al*. also deconstructed a bone-plating procedure into 5 individual tasks[20]. Their results showed no benefit for the blocked group over the random group in skill acquisition for 3 of the 5 tasks.

Skill learning as measured by the difference between the retention score and the pretest score revealed that for the FLS scores, there was a significant difference between the three groups for the peg transfer, ligating loop, IC suturing tasks, and the overall score. The only task that did not show skill learning was the pattern cutting task, where there was no statistical difference seen between the blocked, random, and control groups. Several reasons could account for this. Firstly, the scissors that we used were of low quality and fairly difficult to open and close. Half way through the study we did acquire a better pair but elected not to use them in order to keep the results consistent. Secondly, the penalty score for cutting outside the target area was large (75 points). Even when the subjects would get faster at the task, the penalty for not being accurate would still give many of them a final score of zero.

When using the hand-motion analysis to assess skill learning, we found that for the number of movements a significant difference only existed for the IC suturing task. For the total path-length travelled a significant difference occurred for the ligating loop and IC suturing tasks. For the total time required to complete the task, a significant difference was seen for both the ligating loop and the IC suturing tasks. No difference was seen between the blocked and random groups for any task regardless of the assessment method used. The fact that the same tasks that show a difference using the FLS scores don't necessarily show a difference when using the hand-motion assessment tells us that despite Xeroulis *et al.* showing a high correlation between FLS and hand-motion[38], they obviously measure different attributes of performance.

The contextual interference effect has been found to be fairly robust in the motor skill learning literature. Two main theories have been brought forward to help explain the effect. The first theory that tries to explain this is the *action plan reconstruction* hypothesis[19, 80]. This hypothesis states that before a movement occurs, an "action plan" must be prepared. In blocked practice, a previously prepared "action plan" is readily available from trial to trial, but it suffers from lack of attention on trials following initial retrieval from working memory. Random learners forget the movement solution of each task as they process the task requirements of the other. On switching back to the initial task, they are forced to undergo a demanding reconstructive process to re-plan the way in which they perform the task. It is this need to repeatedly plan the movement solution that results in poorer performance during acquisition. Presumably, this additional trial-to-trial preparation used by the random group results in a more resilient memory representation that better supports long-term recall and transferability of skill.

The second theory is the *elaboration-distinctiveness view* that was initially discussed by Battig[64, 81, 82], but brought forward by Shea and colleagues[15, 83]. The basic idea behind this theory is that during random practice, all to-be-learned tasks remain in working memory and affords the learner many opportunities to compare and contrast the tasks. As a result of this continual comparing and contrasting, rich representations of the tasks are developed and more elaborate and more distinctive memories are established. The need to keep the tasks separate during practice is what causes the disadvantage during acquisition. During blocked practice, the repetition of a long series of the same task makes it less important to keep track of which task is which. Therefore the same elaborate processing and memory links are not established. Although more demanding during acquisition, the need to compare and contrast yields superior performance in retention tests. For transfer tests, the argument is essentially that random practice has made learners more adept at identifying the relevant features of the to-be-performed transfer task, providing an advantage on the tasks, despite the fact that they are novel.

In essence, both of the above theories are based around working memory. In the action-plan reconstruction hypothesis, the previous task is “dumped” from working memory before starting the next task and then regenerated the next time the task needs to be performed. Whereas in the elaboration-distinctiveness view, the tasks all reside beside each other at the same time in working memory during practice. Both theories come to the same conclusion by explaining opposite affects of working memory.

Despite these theories being so pervasive in the motor learning literature and the ubiquitous positive results of the contextual interference effect on practice schedules, we did not find the expected difference between training groups. Several reasons can potentially explain the lack of the contextual interference effect on our results: 1) learner fatigue may have played a role 2) massed practice rather than distributed practice was used 3) not enough practice trials occurred 4) our tasks may have been too complex 5) our retention interval may have been too long 6) It may have been beneficial to have the learners practice several trials of each task in a blocked fashion in order to have a better understanding of each task prior to being randomized to either blocked or random training schedules. We will discuss each of these points in more detail in the following sections.

It has been suggested that training sessions should be limited to approximately one hour, allowing for sufficient concentration to sustain active efforts to improve performance[52]. Training during our study took approximately four hours to complete. Although a short 15-minute break was incorporated into the training schedule, the overall duration of practice exceeded the recommended limit. Trainee fatigue could have contributed to the individual learners not reaching their true potential upon completion of the training session. Ultimately, this also may have contributed to lower retention and transfer test results. Although this would have affected both the blocked and random groups the same since they both practiced the same overall number of times.

We may have been able to overcome learner fatigue by offering the training sessions in a distributed fashion over several days. In a randomized, controlled trial Moulton *et al*. compared training over a single multi-hour session versus training over multiple distributed sessions for a microvascular anastomosis skill among junior surgical residents[50]. Their results showed that both groups improved in performance following training. However, the distributed group performed significantly better on both the retention test and the transfer test. It is thought that between distributed practice sessions, cognitive preparation and mental rehearsal allow for consolidation of the skills into a relatively permanent retention of learned behavior. These key aspects of the skill being learned are similar to the action-plan reconstruction hypothesis.

In order to accurately assess the impact of training schedules on skill learning during distributed training sessions, an approach that allowed a different task to be practiced during each session for the blocked group, and all tasks to be randomly practiced during each session for the random group would likely be appropriate. Tsutsui *et al*. assessed the effects of blocked versus random training over two distributed sessions for a bimanual upper limb coordination task[84]. During experiment 1, they found no contextual interference effect between the 2 groups on retention studies. However, during this experiment participants in the blocked practice group performed each of the tasks on both days of practice. This goes against most of the literature whereby all practice trials are completed at once, and the task is not practiced again during the acquisition period. The authors rectified this with experiment 2, where they used different participants from the first experiment. During experiment 2 the sessions were distributed over 3 days to allow for each of the three tasks to be performed on one day only for the blocked group. The random group again performed all three tasks on all 3 days. This format allowed for the classic contextual interference effect where the blocked group obtained better scores immediately following practice, but the random group outperformed the blocked group during retention testing. This result lends credence to the possibility of assessing distributed training sessions on the effect of blocked versus random practice schedules.

Another interesting finding from Tsutusi *et al*. is the fact that it took 15 trials before a difference in acquisition was seen between the blocked and random groups[84]. This is contrast to what is normally seen in the literature where very early on during practice, a benefit is seen in the blocked group. None of the tasks in our study were performed fifteen times. Similar to the study performed by Tsutusi *et al*.[84], our study used motor tasks that were more complex than those used in most of the contextual interference literature. Had each of the tasks been performed more during our practice sessions, a greater difference might have been seen between the blocked and random groups for all four tasks. These results indicate that for certain skills, the influence of training schedules may lie in the second stage of learning as portrayed by Fitts and Posner[85]. The first stage is the cognitive stage where the learner intellectualizes the task. A thorough explanation and demonstration of the task is required and performance is erratic and performed in distinct steps. With practice and feedback, the learner can reach the integrative stage. This second stage is characterized by deliberate practice and feedback. The learner still needs to consciously think about what to do and how to move. The goal is to comprehend and perform the mechanics more fluidly and with fewer interruptions. The third stage is the autonomous stage. During the final stage practice gradually results in smooth, fluid performance where the learner no longer needs to think about how to execute the task and can concentrate on other aspects of the procedure. The task is able to be completed with speed, efficiency, and precision. The results from Tsutsui *et al*. and our results suggest that for some skills, practice schedules do not influence the processing operations that are involved in figuring out what to do, but instead may manifest more during the second and third stages of learning.

According to the research done on proficiency targets, the rate that different learners reach the second stage of learning will be highly variable. Therefore setting a set number of trials to perform prior to entering a random practice schedule is difficult. Feldman *et al*. had sixteen medical students perform the peg transfer FLS task[45]. They defined the "learning rate" as the number of trials required to get to 90% of the plateau, the theoretical best score achievable. Their results showed that for the group as a whole, 5.9 trials were required to reach 90% of the plateau. However, in the subgroup of students interested in a surgical career, the learning rate was 4.6 trials versus 8.2 trials for those not interested in a surgical career. This result likely indicates that for the peg transfer task, an average of 6 trials is required before the learner enters the second stage of learning. Based on the performance curves for both the FLS scores and the hand-motion analysis, it appears as though the participants in our study reached the second stage of learning for the peg transfer and intracorporeal suturing based on the relative plateau that occurs during those two tasks. Ongoing improvement is still occurring during the final trails of the pattern cutting and ligating loop tasks, indicating that more practice trials are required for those tasks in order to successfully enter the second stage of learning before the contextual interference effect can be allowed to take place.

Longer acquisition phases allow for greater influence of the contextual interference effect. Using a rapid force production task, Shea *et al.* performed experiments looking at the impact of manipulating contextual interference and varying the number of acquisition trials on skill retention[86]. They found that when the acquisition phase consisted of 50 trials, the blocked group outperformed the random group following the completion of acquisition with no difference during retention testing. However, when the acquisition phase allowed for 400 trials the blocked group still outperformed the random group following acquisition, but this time the random group did better during retention testing. The longer acquisition phase is postulated to be the cause for this effect, with more emphasis and demand on working memory for the random group during the prolonged acquisition phase. Participants in our current study performed between 6 and 14 trials depending on the task. Our study uses more complex tasks and therefore a direct comparison cannot be made between the results of the above study and ours. However, it does appear that the number of practice trials has a correlation with the ability to see the effect of the contextual interference.

In addition to the duration of the acquisition phase interacting with the contextual interference, it has been suggested that task complexity has an effect as well. Albaret and Thon studied the effects of task complexity on contextual interference in a drawing task[87]. Task complexity was defined by the number of segments in each pattern. Their results indicated that the effects of contextual interference were only present on the simpler tasks. When completing a more complex task there was no difference seen between the blocked and random group for either retention tests or transfer tests. Current theoretical explanations of the effects of practice schedule during training on motor learning consider that a greater cognitive effort is needed in the random practice compared to the blocked practice condition. Albaret and Thon explained their results by stating that if movements are complex then subjects would have difficulty to maintain all movement related information in working memory from one trial to the next. This would allow for more complex processing strategies even for the blocked group, and would therefore hide the beneficial effect of random practice. Using the action-plan reconstruction hypothesis, the blocked group is able to keep the steps and knowledge of a particular task in working memory. However, if the task is sufficiently complex then this working memory is allowed to degrade, thereby offering the same benefit as the memory-dumping effects of random training, and thus concealing the beneficial effect of random practice.

Task complexity can be broken down into two main categories: functional difficulty and nominal difficulty[88]. Functional difficulty refers to difficulty resulting from the individual performing the task and the environment. For example, the same outdoor task can have a different difficulty depending on whether it is windy or not windy. Nominal difficulty refers to difficulty due to the characteristics of the task only. An example of this is throwing a beanbag into a basket that is close versus a basket that is far away. Random training increases the functional difficulty of a task. However, if the task is already sufficiently difficult (nominal difficulty) that it is already maximizing the efforts of the trainee, then maximum capacity is reached and random training is not adding anything[89]. It is interesting that in our study, although not statistically significant, the easier tasks of peg transfer, pattern cutting, and ligating loop showed better post-test scores for the random group. Only the harder task of intracorporeal suturing showed better post-test scores for the blocked group. Equally surprising is the fact that in arguably the most complex task (IC suturing), the random group showed no degradation in skill on retention tests whereas the blocked group did. Similarly, one of the least complex tasks (pattern cutting) showed no degradation in skill for the blocked group despite the random group showing skill loss.

The effects of practice conditions in motor learning within the constraints of skill level and task difficulty were brought together by Guadagnoli and Lee into the *Challenge Point Framework*[90]. This framework suggests that task difficulties create a learning potential whose function differs according to the level of the performer, the complexity of the task, and the training environment (see figure 42).

This results in a performance-learning paradox where the optimal challenge point for learning does not coincide with the optimal challenge schedule for immediate (acquisition) performance (see figure 43). For a simple task a basic action plan may be developed within the first couple attempts, and further refinement of the skill will be dependent on the extent to which the learner is challenged by the practice conditions. A more complex task may require considerably more time to learn and would require more effort and information processing activities on the part of the learner. It is therefore possible that introducing additional demands for the learner during this process could in fact be detrimental, rather than beneficial since the additional demands may be competing for a limited amount of processing capacity[90]. This same relationship may exist for experienced versus novice performers. When a performer is in a later stage of learning, the systems ability to process information improves, and thus the learner can and should handle more demanding acquisition protocols (see figure 44). One question that remains unanswered is whether learning complex skills late in practice behaves similar to learning simple skills early in practice?

Figure 42. The relationship between nominal task difficulty and expected performance as a function of the individual's skill level. During practice, predicted success becomes a decreasing function of nominal task difficulty. (Used with permission by Taylor & Francis)[90]

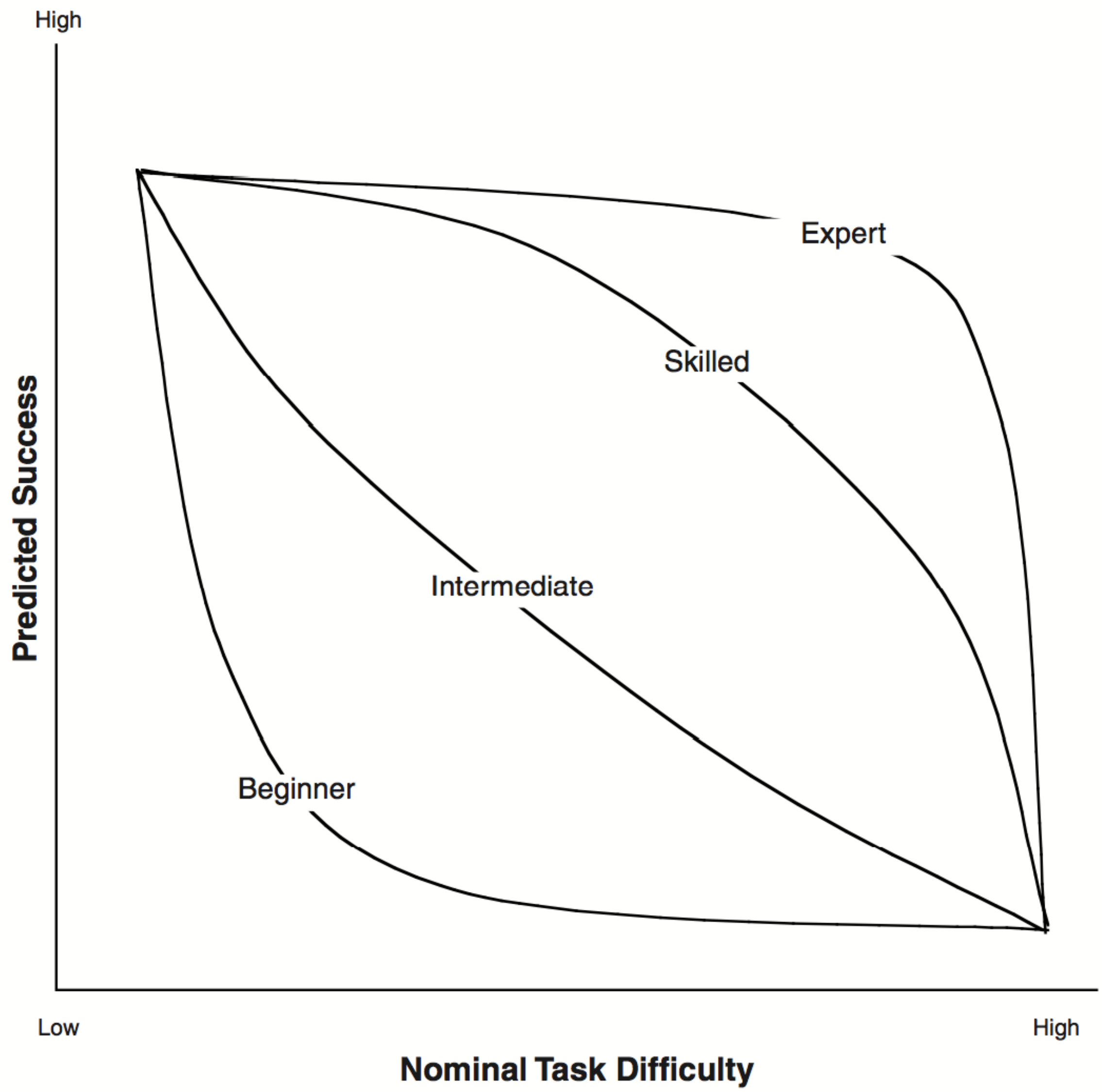


Figure 43. Optimal challenge points for learning (retention) related to different skill levels, functional task difficulty, and potential available information to be learned arising from action, with optimal challenge points resulting earlier (lower functional difficulty) for less experienced performers. (Used with permission by Taylor & Francis)[90]

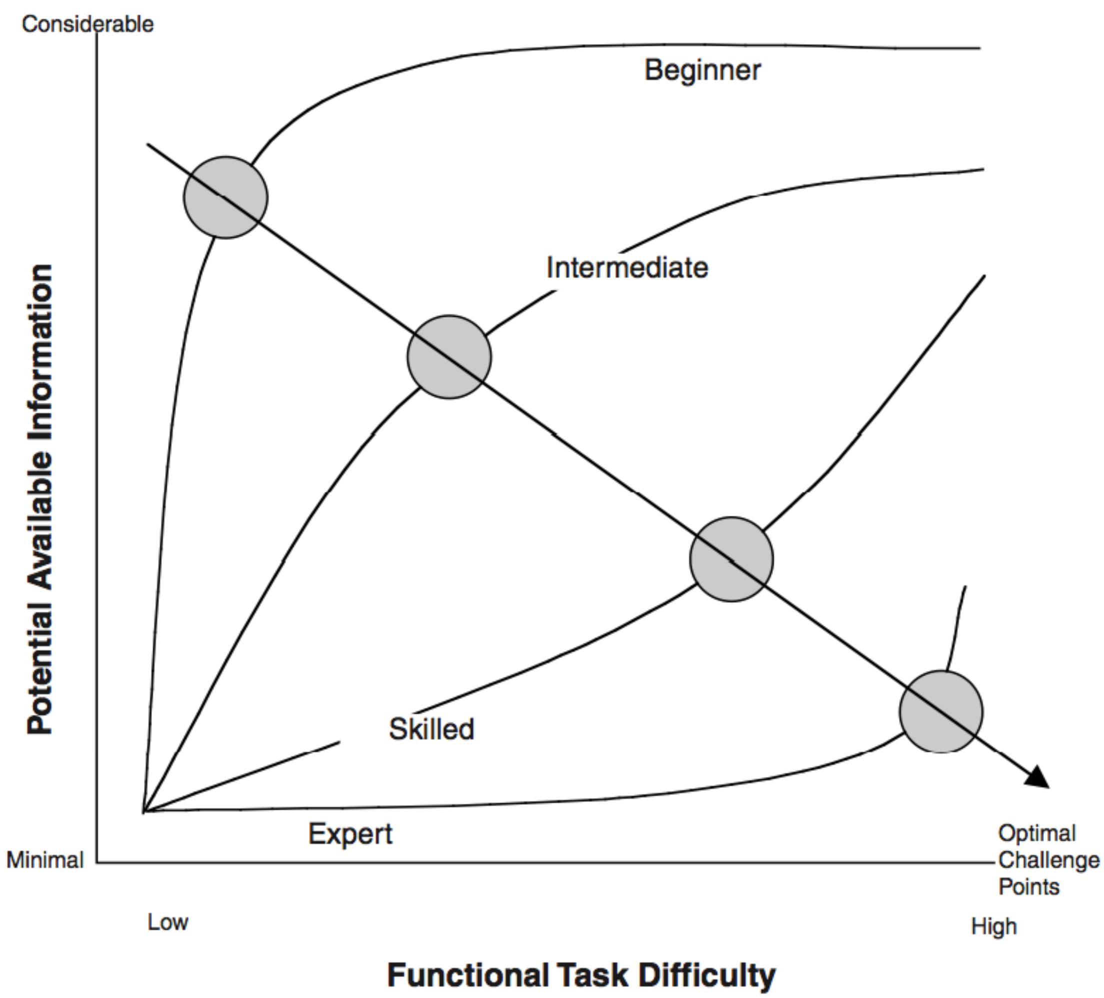
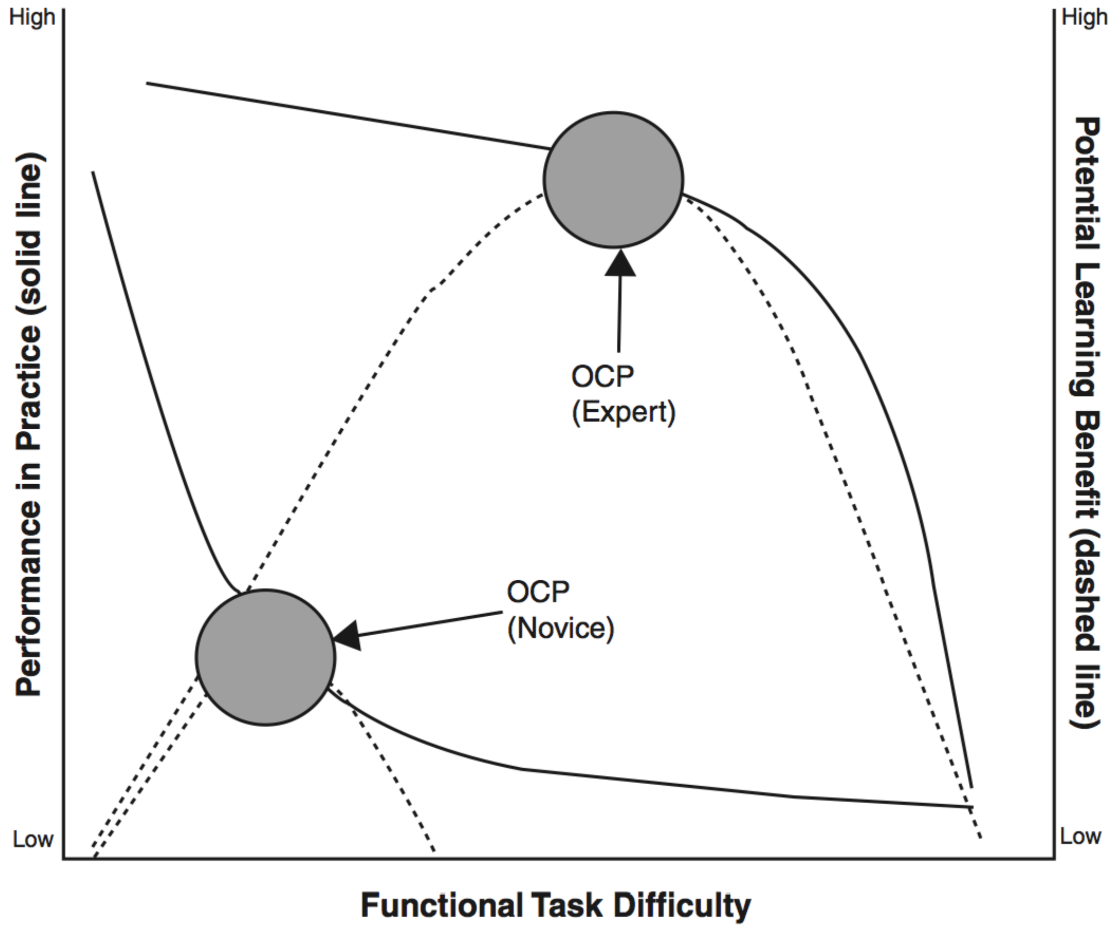


Figure 44. The relation between learning curves, performance curves, and the optimal challenge point (OCP) related to 2 performers of different skill levels. (Used with permission by Taylor & Francis)[90]



Overall, the above theories would suggest that the contextual interference provided by different practice schedules mainly affects the learner during the second stage of learning. More complex tasks also appear to require (at least initially) increased practice in an environment that does not place increased demands on the learner. Therefore it seems likely that the more complex a task is, the more basic (blocked) practice should take place prior to introducing any additional constraints on the learner such as the contextual interference offered by random practice schedules. Early practice should focus on getting it right. Only once a learner knows what to do, will the true benefit from a random practice schedule appear[91].

The effect of nominal task difficulty interfering with the contextual interference has been further highlighted by Jarus and Gutman[92]. They performed a study where children practiced a beanbag toss under simple or complex task conditions. The simple condition involved tossing beanbags of different weight into targets of a constant order. The complex task involved changing the weight of the beanbags and the order of the targets. In their study, practicing the complex task under blocked conditions resulted in better performance over the random group during the acquisition, retention, and transfer tasks. It was hypothesized that the complex task caused too much impedance on the random practice group and thereby hindered learning. Another possibility put forth by the authors as to why Battig's predictions weren't fully supported is due to the relatively low number of practice trials (30 practice trials). However, in the simple task, practicing under random conditions showed no difference compared to the blocked group during the acquisition phase, but better performance during the retention and transfer tasks. Similarly, functional task difficulty using various tennis ground strokes among low-skilled and high-skilled learners was assessed by Hebert *et al.*[93]. Their study showed that participants defined as low-skilled learned better with a blocked schedule while the high-skilled learned better with a random training schedule. The subjects in our study were all novices in laparoscopic surgery, putting them in the low-skilled category. During the early stage of their careers, this would imply that blocked training would be the most beneficial to learning. As they gain more laparoscopic skills, and progress to the high-skilled category, they would also benefit more from a random training schedule.

In addition to the practice schedule modifying the contextual interference, it has also been suggested that the amount of similarity between tasks also contributes to the amount of contextual interference. Boutin and Blandin performed a study where participants practiced a set of tasks that were all either similar or dissimilar in either blocked or random order[94]. The tasks involved depressing microswitches with various timing goals. The similar group always used the same microswitch, but varied the timing. The dissimilar group alternated between three different microswitches as well as the timing. Interestingly, they found that the contextual interference effect of random practice only benefited retention scores when the tasks were similar. The results were explained using the challenge point framework. Practicing the dissimilar tasks in random order created too much interference and overwhelmed the learning capacity of the learner. In our present study the tasks can be viewed as being both complex and dissimilar to each other, which may further explain why the contextual interference effect was not seen.

An argument could be made for overtraining as to the reason why there was no difference seen between the blocked and random groups. Training beyond the passing level may be beneficial and has been shown to improve skill retention by Scott *et al*.[14]. However, the amount of trials performed in that study were 57 for the peg transfer task, 18 for the pattern cutting task, 8 for the ligating loop task, and 28 for the intracorporeal suturing task. Using this amount of overtraining, the subjects achieved a 100% pass rate based on the normalized score of 270 required to pass the technical skills component of the FLS curriculum[68]. Participants in our study performed 14, 8, 6, and 12 trials for the peg transfer, pattern cutting, ligating loop, and intracorporeal suturing tasks respectively. Other than the ligating loop task, it can be seen that participants in our study performed much fewer trials, and as such overtraining likely did not occur. In fact, we made the number of trials in our study low in order to prevent overtraining. However, the numbers we chose were likely too low, which may have inhibited the contextual interference effect.

Because all five tasks are given approximately equal weight in comprising the overall pass score and since our participants did not perform the extracorporeal suturing task, a score of approximately 220 would be required to pass the four tasks in our study. None of the participants in either of the blocked, random, or control groups would have achieved a passing score prior to practicing. However, all participants in the blocked and random groups would have passed immediately following training. During retention testing, eight of the thirteen in the blocked group would have passed, ten of the fourteen in the random group would have passed, and none of the participants in the control group would have passed. Based on the far fewer trials of each task practiced and the fact that many of the participants did not attain a passing score during retention testing, we can assume that an element of overtraining was not responsible for the absence of a difference between our training groups.

Similar to the retention tests, the transfer test in our study did not reveal a difference between the groups using either the FLS scoring ($p = 0.49$) or the hand-motion analysis scores (number of movements, p = 0.41; path-length travelled, p = 0.27; time to completion, p = 0.45). Relative to the four tasks used during the practice sessions, the transfer task can be seen as being less difficult. Shea and Morgan found that performance on transfer tasks was better amongst the random group when the task was more complex[15]. However, the tasks practiced during the acquisition phase of their study were relatively simple. As has already been discussed above, complex skills likely have different learning principles and variable effects than simple skills and therefore a generalization to our study may not be possible[95].

We were unable to show a difference between the blocked and random training schedules. Using what has been learned from this current study, future studies should contain larger group sizes in order to detect a smaller effect size than was postulated for our current study (30 participants per group would be able to detect an effect size of 0.7). Additionally, distributed practice should take place so that learner fatigue is minimized. Multiple sessions should be organized such that the blocked group practices a separate task during each session and the random group practices all tasks during each session. Performing 16 tasks per session would allow the blocked group to complete all four tasks in 4 sessions and allow the random group to practice each task 4 times during each session. Practicing each task 16 times would hopefully also allow for enough trials to show the contextual interference effect. Since the FLS tasks are complex and our target study population is inexperienced, a short interval of blocked training prior to being randomized into either blocked or random practice schedules would be beneficial. A retention interval of shorter than 5-6 weeks may also be beneficial. Although several studies have shown that motor skills can be retained for several months, these are not for complex skills such as those in our study. An initial retention interval of 3-4 weeks may be better suited to our study. A second retention interval at a total of 6-8 weeks could also be looked at to see if with the other above modifications, a longer retention period is possible. Performing such a large study would likely require collaboration from multiple institutions as well as very careful scheduling in order to integrate with the busy clinical schedule of surgical trainees.

A. Limitations of this Study

The present study had some limitations. Firstly, we included junior surgical residents from specialties that do not perform laparoscopic surgery. While we believe that in the early stages of surgical residency laparoscopic skills likely would not have been developed, there is a possibility that the participants from the "laparoscopic" specialties may have entered into this study already some point along their learning curves, thereby changing the overall effect of learning that took place during the study period. Additionally, the laparoscopic specialties may have received additional practice between the immediate post-test and the retention and transfer tests. We also included senior medical students with a self-declared interest in a surgical specialty. Again, while a senior medical student and a junior surgical resident likely would not have a significant difference in surgical skill level, the results may have been different with a study population made up entirely of residents, since they may be a little further along in the low-skilled to high-skilled continuum. Secondly, while we did perform formal sample size calculations, the effect size that we used may have been too large. An effect size of 1.2 is likely able to discriminate between training and no training, but may be larger than the effect size between two different training groups. We were able to show that both training schedules were beneficial over the control group. However, if the true effect size comparing one intervention over the other were smaller, we would have required a larger sample size to show a difference between the blocked and random practice groups. Thirdly, given the busy nature of a junior resident and a senior medical student's clinical schedule, setting a specific time delay between the initial training session and the retention and transfer test was not practical. Despite the fact that there was no difference in the average time delay between the groups, there was a fair amount of variation in the time delay between individual subjects. Individuals may retain or forget the surgical skills that they learned at different rates, potentially affecting the results of the present study. Despite these limitations, all groups and all tasks likely would have been equally affected.

Chapter 9. Conclusion

The fundamentals of Laparoscopic Surgery scoring system and computer-based hand-motion analysis metrics are both well-validated means of assessing surgical skill. Using these two assessment methods among senior medical students and junior surgical residents we found that skill learning, as measured by retention scores shows a positive benefit for practicing using either the blocked or random format. However, we were unable to show a difference between the blocked group and the random group during acquisition, retention, or transfer tests.

Training to proficiency goals appears to have a lasting effect on retention in the literature. The optimal training schedule to obtain these proficiency goals is still undetermined. Currently the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and the FLS program endorse the use of blocked practice schedules. Our study does not support blocked training over random training schedules for the learning, retention, and transfer of complex tasks. Therefore both training schedules are options available to be used by the learner.

Our results do not support the use of one training schedule over another. However, the psychomotor learning literature provides theory and evidence in support of the benefits of random practice schedules for the learning and retention of motor skills, therefore making it a reasonable option. Future studies should focus on including more participants in order to detect a smaller effect size, incorporating more practice trials, using a distributed practice format, and allowing the learners to perform several trials of each task in a blocked fashion prior to being randomized into a blocked or random practice group. It is possible that for complex surgical skills a combination of blocked training followed by random training schedules will allow for the most learning.

Our study adds to the current literature suggesting that complex motor tasks have different learning patterns from that of simple motor tasks. More research is required before firm surgical training curriculum suggestions can be made.

Bibliography

1. Halsted, W.S., *The training of the surgeon.* Bull Johns Hop Hosp, 1904. **xv**: p. 267-275.
2. Philibert, I., P. Friedmann, and W.T. Williams, *New requirements for resident duty hours.* JAMA, 2002. **288**(9): p. 1112-4.
3. Gates, E.A., *New surgical procedures: can our patients benefit while we learn?* Am J Obstet Gynecol., 1997. **176**(6): p. 1293-8; discussion 1298-9.
4. Lossing, A.G., et al., *A technical-skills course for 1st-year residents in general surgery: a descriptive study.* Can J Surg, 1992. **35**(5): p. 536-40.
5. Reznick, R.K., *Teaching and testing technical skills.* Am J Surg, 1993. **165**(3): p. 358-61.
6. Bell, R.H., *Surgical council on resident education: a new organization devoted to graduate surgical education.* J Am Coll Surg, 2007. **204**(3): p. 341-6.
7. Sachdeva, A.K., C.A. Pellegrini, and K.A. Johnson, *Support for simulation-based surgical education through American College of Surgeons--accredited education institutes.* World J Surg, 2008. **32**(2): p. 196-207.
8. Soper, N.J. and G.M. Fried, *The fundamentals of laparoscopic surgery: its time has come.* Bull Am Coll Surg, 2008. **93**(9): p. 30-2.
9. Peters, J.H., et al., *Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery.* Surgery, 2004. **135**(1): p. 21-27.
10. Derossis, A.M., et al., *Development of a model for training and evaluation of laparoscopic skills.* Am J Surg, 1998. **175**(6): p. 482-7.
11. Fried, G.M. and L.S. Feldman, *Objective assessment of technical performance.* World J Surg, 2008. **32**(2): p. 156-60.
12. Fried, G.M., et al., *Proving the Value of Simulation in Laparoscopic Surgery.* Annals of Surgery, 2004. **240**(3): p. 518-528.
13. Scott, D.J. *Fundamentals of Laparoscopic Surgery: Technical skills proficiency-based curriculum*. 2006 [cited 2012 April 3]; Available from: http://www.flsprogram.org/wp-content/uploads/2010/08/Technical-Skills-Proficiency-Bases-Training-Curriculum.pdf.
14. Scott, D.J., et al., *Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training.* Surg Endosc, 2008. **22**(8): p. 1887-93.
15. Shea, J.B. and R.L. Morgan, *Contextual interference effects on the acquisition, retention, and transfer of a motor skill.* Journal of Experimental Psychology: Human Learning and Memory, 1979. **5**(2): p. 179-187.
16. Dubrowski, A. and D. Backstein, *The contributions of kinesiology to surgical education.* J Bone Joint Surg Am, 2004. **86-A**(12): p. 2778-81.
17. Dubrowski, A., et al., *The influence of practice schedules in the learning of a complex bone-plating surgical task.* Am J Surg, 2005. **190**(3): p. 359-63.
18. Dubrowski, A., *Performance vs. learning curves: what is motor learning and how is it measured?* Surg Endosc, 2005. **19**(9): p. 1290.
19. Lee, T.D. and R.A. Magill, *Can forgetting facilitate skill acquisition?*, in *Differing Perspectives in Motor Learning, Memory, and Control*, D. Goodman, R.B.

Wilberg, and I.M. Franks, Editors. 1985, Elsevier: Amsterdam, the Netherlands. p. 3-22.
20. Brydges, R., et al., *Application of motor learning principles to complex surgical tasks: searching for the optimal practice schedule.* J Mot Behav, 2007. **39**(1): p. 40-8.
21. Bridges, M. and D.L. Diamond, *The financial impact of teaching surgical residents in the operating room.* Am J Surg, 1999. **177**(1): p. 28-32.
22. Hawasli, A., et al., *Laparoscopic training in residency program.* J Laparoendosc Surg, 1996. **6**(3): p. 171-4.
23. Beyer, L., et al., *Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study.* Am J Surg, 2011. **202**(3): p. 265-72.
24. Grantcharov, T.P., et al., *Randomized clinical trial of virtual reality simulation for laparoscopic skills training.* Br J Surg, 2004. **91**(2): p. 146-50.
25. Park, J., et al., *Randomized controlled trial of virtual reality simulator training: transfer to live patients.* Am J Surg, 2007. **194**(2): p. 205-11.
26. Scott, D.J., et al., *Laparoscopic training on bench models: better and more cost effective than operating room experience?* J Am Coll Surg, 2000. **191**(3): p. 272-83.
27. Seymour, N.E., et al., *Virtual reality training improves operating room performance: results of a randomized, double-blinded study.* Ann Surg, 2002. **236**(4): p. 458-63; discussion 463-4.
28. Sroka, G., et al., *Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial.* Am J Surg, 2010. **199**(1): p. 115-20.
29. Derossis, A.M., et al., *The effect of practice on performance in a laparoscopic simulator.* Surg Endosc, 1998. **12**(9): p. 1117-20.
30. Martin, J.A., et al., *Objective structured assessment of technical skill (OSATS) for surgical residents.* Br J Surg, 1997. **84**(2): p. 273-8.
31. Regehr, G., et al., *Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination.* Acad Med, 1998. **73**(9): p. 993-7.
32. Reznick, R., et al., *Testing technical skill via an innovative "bench station" examination.* Am J Surg, 1997. **173**(3): p. 226-30.
33. van der Vleuten, C.P.M. and D.B. Swanson, *Assessment of clinical skills with standardized patients: State of the art.* Teach Learn Med, 1990. **2**(2): p. 58-76.
34. Taffinder, N., et al., *Can a computer measure surgical precision? Reliability, validity and feasibility of the ICSAD.* Surg Endosc, 1999. **13**(Suppl 1): p. 81.
35. Datta, V., et al., *Motion Analysis in the Assessment of Surgical Skill.* Computer Methods in Biomechanics and Biomedical Engineering, 2001. **4**(6): p. 515-523.
36. Datta, V., et al., *The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model.* J Am Coll Surg, 2001. **193**(5): p. 479-85.
37. Datta, V., et al., *The relationship between motion analysis and surgical technical assessments.* Am J Surg, 2002. **184**(1): p. 70-3.

38. Xeroulis, G., A. Dubrowski, and K. Leslie, *Simulation in laparoscopic surgery: a concurrent validity study for FLS.* Surg Endosc, 2009. **23**(1): p. 161-5.
39. Anastakis, D.J., et al., *Assessment of technical skills transfer from the bench training model to the human model.* Am J Surg, 1999. **177**(2): p. 167-70.
40. Gurusamy, K.S., et al., *Virtual reality training for surgical trainees in laparoscopic surgery.* Cochrane Database Syst Rev, 2009(1): p. CD006575.
41. Fowler, D.L., *Enabling, implementing, and validating training methods in laparoscopic surgery.* World J Surg, 2010. **34**(4): p. 621-4.
42. Hogle, N.J., et al., *Validation of laparoscopic surgical skills training outside the operating room: a long road.* Surg Endosc, 2009. **23**(7): p. 1476-82.
43. Scott, D.J., et al., *Laparoscopic skills training.* Am J Surg, 2001. **182**(2): p. 137-42.
44. Brunner, W.C., et al., *Laparoscopic virtual reality training: are 30 repetitions enough?* J Surg Res, 2004. **122**(2): p. 150-6.
45. Feldman, L.S., et al., *A method to characterize the learning curve for performance of a fundamental laparoscopic simulator task: defining "learning plateau" and "learning rate".* Surgery, 2009. **146**(2): p. 381-6.
46. Ritter, E.M. and D.J. Scott, *Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery.* Surg Innov, 2007. **14**(2): p. 107-12.
47. Gallagher, A.G., et al., *Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training.* Ann Surg, 2005. **241**(2): p. 364-72.
48. Stefanidis, D., *Optimal acquisition and assessment of proficiency on simulators in surgery.* Surg Clin North Am, 2010. **90**(3): p. 475-89.
49. Ericsson, K.A., R.T. Krampe, and C. Tesch-Romer, *The role of deliberate practice in the acquisition of expert performance.* Psychol Rev, 1993(100): p. 363-406.
50. Moulton, C.A., et al., *Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial.* Ann Surg, 2006. **244**(3): p. 400-9.
51. Luft, A.R. and M.M. Buitrago, *Stages of motor skill learning.* Mol Neurobiol., 2005. **32**(3): p. 205-16.
52. Kolozsvari, N.O., et al., *Sim one, do one, teach one: considerations in designing training curricula for surgical simulation.* J Surg Educ, 2011. **68**(5): p. 421-7.
53. Donovan, J. and D. Radosevich, *A meta-analytic review of the distribution of practice effect: Now you see it, now you don't.* Journal Appl Psychol, 1999. **84**(5): p. 795-805.
54. *Retention*, in *Merriam-Webster Online Dictionary*. p. Retention.
55. Schmidt, R.A. and R.A. Bjork, *New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training.* Psychological Science, 1992. **3**(4): p. 207-217.
56. Arthur, W., et al., *Factors That Influence Skill Decay and Retention: A Quantitative Review and Analysis.* Human Performance, 1998. **11**(1): p. 57-101.
57. Edelman, D.A., M.A. Mattos, and D.L. Bouwman, *FLS skill retention (learning) in first year surgery residents.* J Surg Res, 2010. **163**(1): p. 24-8.

58. Castellvi, A.O., et al., *Maintaining proficiency after fundamentals of laparoscopic surgery training: a 1-year analysis of skill retention for surgery residents.* Surgery, 2009. **146**(2): p. 387-93.
59. Mashaud, L.B., et al., *Two-year skill retention and certification exam performance after fundamentals of laparoscopic skills training and proficiency maintenance.* Surgery, 2010. **148**(2): p. 194-201.
60. Stefanidis, D., et al., *Skill retention following proficiency-based laparoscopic simulator training.* Surgery, 2005. **138**(2): p. 165-70.
61. Stefanidis, D., et al., *Proficiency maintenance: impact of ongoing simulator training on laparoscopic skill retention.* J Am Coll Surg, 2006. **202**(4): p. 599-603.
62. Shadmehr, R. and T. Brashers-Krug, *Functional stages in the formation of human long-term motor memory.* J Neurosci, 1997. **17**(1): p. 409-19.
63. Hikosaka, O., et al., *Long-term retention of motor skill in macaque monkeys and humans.* Exp Brain Res, 2002. **147**(4): p. 494-504.
64. Battig, W.F., *Facilitation and interference*, in *Acquisition of Skill*, E.A. Bilodeau, Editor 1966, Academic Press: New York. p. 215-244.
65. Matsumoto, E.D., et al., *The effect of bench model fidelity on endourological skills: a randomized controlled study.* J Urol, 2002. **167**(3): p. 1243-7.
66. Cohen, J., *Statistical power analysis for the behavioral sciences*1977, New York: Academic Press.
67. Grober, E.D., et al., *The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures.* Ann Surg, 2004. **240**(2): p. 374-81.
68. Fraser, S.A., et al., *Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system.* Surg Endosc, 2003. **17**(6): p. 964-7.
69. McCluney, A.L., et al., *Laparoscopic Cannulation: Use of the SAGES Learning Center to Validate a Potentially New FLS Task, Poster session presented at: 2007 Scientific Session of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES), Las Vegas, Nevada, USA, 18-22 April 2007. Abstracts.* Surg Endosc, 2007. **21 Suppl 1**: p. S340-351.
70. Fried, G.M., *FLS assessment of competency using simulated laparoscopic tasks.* J Gastrointest Surg, 2008. **12**(2): p. 210-2.
71. Brydges, R., et al., *Construct validity of computer-assisted assessment: quantification of movement processes during a vascular anastomosis on a live porcine model.* Am J Surg, 2007. **193**(4): p. 523-9.
72. Dubrowski, A., et al., *Quantification of process measures in laparoscopic suturing.* Surg Endosc, 2006. **20**(12): p. 1862-6.
73. Dubrowski, A., et al., *Quantification of motion characteristics and forces applied to tissues during suturing.* Am J Surg, 2005. **190**(1): p. 131-6.
74. Grober, E.D., et al., *Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity.* Microsurgery, 2003. **23**(4): p. 317-22.
75. Moorthy, K., et al., *Motion analysis in the training and assessment of minimally invasive surgery.* Minim Invasive Ther Allied Technol, 2003. **12**(3): p. 137-42.

76. Vassiliou, M.C., et al., *Cannulation: A Potential Addition to the FLS Program*, in *Poster session presented at: Society of American Gastrointestinal and Endoscopic Surgeons*2012 March 7-10: San Diego, CA.
77. Rosenbaum, D.A., *Human motor control*1991: Academic Press.
78. Bann, S.D., M.S. Khan, and A.W. Darzi, *Measurement of surgical dexterity using motion analysis of simple bench tasks.* World J Surg, 2003. **27**(4): p. 390-4.
79. Smith, S.G., et al., *Motion analysis.* Surg Endosc, 2002. **16**(4): p. 640-5.
80. Lee, T.D. and R.A. Magill, *The locus of contextual interference in motor-skill acquisition.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 1983(9): p. 730-746.
81. Battig, W.F., *Intratask interference as a source of facilitation in transfer and retention*, in *Topics in learning and performance*, R.F. Thompson and J.F. Voss, Editors. 1972, Academic Press: New York. p. 131-159.
82. Battig, W.F., *The flexibility of human memory*, in *Levels of processing in human memory*, L.S. Cermak and F.I.M. Craik, Editors. 1979, Erlbaum: Hillsdale, NJ. p. 23-44.
83. Shea, J.B. and S.T. Zimny, *Context Effects in Memory and Learning Movement Information*, 1983. p. 345-366.
84. Tsutsui, S., T.D. Lee, and N.J. Hodges, *Contextual interference in learning new patterns of bimanual coordination.* J Mot Behav, 1998. **30**(2): p. 151-7.
85. Fitts, P.M. and M.I. Posner, *Human performance*1967, Belmont, CA: Brooks/Cole Pub. Co.
86. Shea, C.H., R. Kohl, and C. Indermill, *Contextual interference: Contributions of practice.* Acta Psychologica, 1990. **73**(2): p. 145-157.
87. Albaret, J.M. and B. Thon, *Differential effects of task complexity on contextual interference in a drawing task.* Acta Psychol (Amst), 1998. **100**(1-2): p. 9-24.
88. Swinnen, S.P., et al., *The effect of movement speed on upper-limb coupling strength.* Human Movement Science, 1992. **11**(5): p. 615-636.
89. Marteniuk, R.G., *Information Processing in Motor Skills*1976: Holt, Rinehart and Winston.
90. Guadagnoli, M.A. and T.D. Lee, *Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning.* J Mot Behav, 2004. **36**(2): p. 212-24.
91. Magill, R.A. and K.G. Hall, *A review of the contextual interference effect in motor skill acquisition.* Human Movement Science, 1990. **9**(3–5): p. 241-289.
92. Jarus, T. and T. Gutman, *Effects of cognitive processes and task complexity on acquisition, retention, and transfer of motor skills.* Can J Occup Ther., 2001. **68**(5): p. 280-9.
93. Hebert, E.P., D. Landin, and M.A. Solmon, *Practice schedule effects on the performance and learning of low- and high-skilled students: an applied study.* Res Q Exerc Sport., 1996. **67**(1): p. 52-8.
94. Boutin, A. and Y. Blandin, *On the cognitive processes underlying contextual interference: Contributions of practice schedule, task similarity and amount of practice.* Hum Mov Sci, 2010. **29**(6): p. 910-20.
95. Wulf, G. and C.H. Shea, *Principles derived from the study of simple skills do not generalize to complex skill learning.* Psychon Bull Rev., 2002. **9**(2): p. 185-211.

Appendix 1. Questionnaire Given to Potential Study Participants

Participant Questionnaire

Identification number: ______________________________

1. Are you a medical student? Y / N

   If yes, please list medical school year

2. Are you a resident? Y / N

   If yes, please list your program name and post-grad year

3. Age ____

4. Gender F / M

5. What hand are you R / L

6. How much experience do you have with laparoscopic procedures?

   Number of laparoscopic procedures observed ____

   Number of procedures acting as camera-person ____

   I have performed operative procedures Y / N

   If yes, please describe number and nature of performance.

7. Have you practiced on any other laparoscopic box trainers or computer-based simulators before? Y / N

   If yes, please describe nature of performance.

Appendix 2. Form Used to Obtain Informed Consent

RESEARCH PARTICIPANT INFORMATION AND CONSENT FORM

Title of Study: The Effect of Blocked Versus Random Task Practice Schedules on the Acquisition, Retention, and Transfer of Laparoscopic Skills

Principal Investigator
Dr. Jason Park
Address: St. Boniface Hospital, 409 Tache Avenue, 3rd Floor Z-Block, Winnipeg, MB R2H 2A6
Phone number: 237-2574

Co-investigators
Dr. Justin Rivard, Department of Surgery, University of Manitoba, Phone number 803-4222
Dr. Ashley Vergis, Department of Surgery, University of Manitoba, Phone number 237-2574

You are being asked to participate in a research study. Please take your time to review this consent form and discuss any questions you may have with the study staff. You may take your time to make your decision about participating in this study and you may discuss it with your friends, family or (if applicable) your doctor before you make your decision. This consent form may contain words that you do not understand. Please ask the study staff to explain any words or information that you do not clearly understand.

Exclusion Criteria
You must have no prior laparoscopic surgical experience, including simulated training.

Purpose of Study
We want to determine how the ordering of laparoscopic surgical skills training affects the speed of learning, skill retention, and cross-task skill translation.

Study procedures
You are being asked to participate in a laparoscopic skills training course as part of the above study. The course will last up to four hours and will be scheduled at a time that is convenient for you. You will have your laparoscopic performance video-recorded and assessed before and immediately after the course. You will have your laparoscopic performance assessed again four weeks after the course, scheduled again at a time convenient for you.

You can stop participating at any time. However, if you decide to stop participating in the study, we encourage you to talk to the study staff first. If you are interested in the results of the study you may contact the Principal Investigator at the end of the study.

Discomforts
You may be uneasy about having your laparoscopic skills assessed. We therefore emphasize that you will remain anonymous and may also withdraw from the study at any time.

Benefits
You will gain experience in basic laparoscopic surgical skills, which could be useful during your clinical rotations on surgical services. These are skills that residents in general surgery, urology, and obstetrics and gynecology and students who go on to surgical residencies in these specialties may need to perform.

Payment for Participation
You will receive an honoraria consisting of a $100 plus a gift certificate for $12, which you can use to purchase refreshments prior to and during your training sessions.

Confidentiality
Information gathered in this research study may be published or presented in public forums; however your name and other identifying information will not be used or revealed. Only the study investigators will have access to the videotapes. All videotapes will be destroyed 5 years after the study ends. Despite efforts to keep your personal information confidential, absolute confidentiality cannot be guaranteed. Your personal information may be disclosed if required by law.
The University of Manitoba Health Ethics Research Board may review records related to the study for quality assurance purposes.

No information revealing any personal information such as your name will leave the University of Manitoba.

Voluntary Participation/Withdrawal from the Study
Your decision to take part in this study is voluntary. You may refuse to participate or you may withdraw from the study at any time.

Participants who are students or employees of either The University of Manitoba or Health Sciences Centre or individuals associated professionally with any of the investigators can be assured that a decision not to participate will in no way affect any performance evaluation of potential participants.

Questions

You are free to ask any questions that you may have about your treatment and your rights as a research participant. If any questions come up during or after the study or if you have a research-related injury, contact any one of the study investigators: Jason Park (204) 237-2574, Justin Rivard (204) 803-4222, Ashley Vergis (204) 237-2594

For questions about your rights as a research participant, you may contact the University of Manitoba, Bannatyne Campus Research Ethics Board Office at (204) 789-3389.

Do not sign this consent form unless you have had a chance to ask questions and have received satisfactory answers to all of your questions.

Statement of Consent

I have read this consent form. I have had the opportunity to discuss this research study with the study investigators. I have had my questions answered by them in language I understand. The risks and benefits have been explained to me. I believe that I have not been unduly influenced by any study team member to participate in the research study by any statements or implied statements. Any relationship (such as employer, supervisor or family member) I may have with the study team has not affected my decision to participate. I understand that I will be given a copy of this consent form after signing it. I understand that my participation in this study is voluntary and that I may choose to withdraw at any time. I freely agree to participate in this research study.

I understand that information regarding my personal identity will be kept confidential, but that confidentiality is not guaranteed. I authorize the inspection of any of my records that relate to this study by The University of Manitoba Research Ethics Board, for quality assurance purposes.

By signing this consent form, I have not waived any of the legal rights that I have as a participant in a research study.

Participant signature: _______________________ Date:____________

Participant printed name: _______________________

I, the undersigned, have fully explained the relevant details of this research study to the participant and believe that the participant named above has understood and has knowingly given their consent.

Printed Name: ________________________ Date: _____________

Signature: ________________________

Role in the study: ________________________

Relationship (if any) to study team members: _____________________