Theses and Dissertations   Graduate School

2012

DETERMINATION OF EFFECTIVE TRAINING METHODS TO LEARN A LAPAROSCOPIC CAMERA NAVIGATION TASK UNDER STRESSFUL ENVIRONMENTS

Devnath Vasudevan
Virginia Commonwealth University

Determination of Effective Training methods to learn a Laparoscopic Camera Navigation Tasks under Stressful Environments

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

at Virginia Commonwealth University.

By

Devnath Vasudevan

Department of Biomedical Engineering, Virginia Commonwealth University

Director: Dianne Pawluk, Ph.D., Assistant Professor, Department of Biomedical Engineering

Virginia Commonwealth University

Richmond, Virginia, USA

May, 2012

Acknowledgement

Table of Contents

iv

ix

x

List of Tables

List of Figures

xviii

Abstract

DETERMINATION OF EFFECTIVE TRAINING METHODS TO LEARN A LAPAROSCOPIC CAMERA NAVIGATION TASK UNDER STRESSFUL ENVIRONMENTS

By Devnath Vasudevan, M.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science, at Virginia Commonwealth University.

Virginia Commonwealth University, 2012

Major Director: Dr.Dianne Pawluk, Assistant Professor, Department of Biomedical Engineering

Stress in surgical environment is generally very high and can result in performance degradation increasing patient risk .Current Training systems for learning minimally invasive surgical skills do not consider the component of stress in their training model. In this study the focus was on

developing alternative training models that would allow the learner to effectively perform

minimally invasive skill under stress. Two alternate training methods: 1) Training under stress

until high performance levels and 2) training until high performance and low cognitive load are

achieved were considered for this study. The control group consisted of training under no stress

and until high performance levels are achieved. Stressful environments for this study were

simulated using physiologic stressors. The effectiveness of the training was evaluated by a

comparative analysis of the different performance measures across the groups. We determined

that training until automation as the most effective method to perform effectively under stress.

Chapter 1

Introduction

Surgery is a medical field consisting of physical intervention to tissues and organs by employing operative treatment. Based on the extent of the intervention, surgical procedures can be broadly classified into invasive (open), minimally invasive, and non-invasive. Open surgeries are characterized by larger incisions, higher blood loss and longer recuperative time when compared to other types of surgery. Advances in the field of surgery have resulted in novel techniques to perform procedures which decrease these problems, potentially benefiting patients. One such technique is the Minimally Invasive Surgical (MIS) procedure. This procedure is performed by making small incisions in the surgical field through which long instruments and cameras are inserted. Although this technique provides potential benefits to the patient, it also presents some unique problems for the surgeon. These include:

1. A restricted freedom of tool movement. – The tools are inserted via a trocar in the incision site. As a result, movement is limited by the pivot point created by the trocar to 4 degrees-of-freedom (4 DOF).

2. An indirect view of the workspace. – The surgical workspace is viewed via a laparoscopic camera. That view is often non-collocated with the surgical instruments, requiring the surgeon to re-orient their perspective.

3. The presentation of the 3-dimensional workspace to a 2-dimensional video image. – Depth perspective is lost in this transformation.

4. The location of the laparoscopic video monitor. – Off-axis viewing of the laparoscopic image often further complicates the spatial relationship the surgeon has with the operative field.

5. The use of long surgical tools. – This modifies the mechanical relationship between the surgeon and the patient.

6. The reduction or lack of force feedback through the tools. – MIS tools are currently limited in the amount of haptic feedback they provide the surgeon.

These factors combine to place increased cognitive and physical demands on the surgeons, making MIS inherently more difficult to learn and perform than open surgery. This drawback is offset, however, secondary to the patient-specific benefits MIS offers. As a result, MIS is gaining popularity and many new surgical procedures are being performed with this technique.

For a surgical procedure to be successful, surgeons must be relatively error free while multitasking and meeting the various demands of the environment effectively. Expert surgeons achieve this through the knowledge and experience gained by performing many operations over a number of years. But the growing demand for laparoscopic surgeries has resulted in the need for novices to master the required skills within a short time period with a high degree of

2

accuracy. The old method of developing these skills included training under real conditions in the operating room (O.R.). This opportunity is becoming increasingly limited due to time restrictions on residents in the O.R. As a result, there is a current need to develop effective training methods that can help surgeons master these skills outside the O.R. environment. This has become the focus of this and many research studies.

The goal of this study is twofold:

1. Develop a training method that would enable the learner to learn to effectively handle a laparoscope which is a task that is very essential for performing MIS procedures

2. Develop a training method that would enable the learner to effectively handle stress which is prevalent in the OR and a major reason for performance deterioration

A key component in the MIS surgical setup is the laparoscope, a camera with a telescopic shaft lens system that is inserted through narrow openings (ports) in the person's body by the surgeon to view the operating field. The surgeon relies on this visual information to manipulate the surgical tools in order to perform the surgery. It is therefore very important that the laparoscope is handled effectively to ensure success of the surgery.

A fundamental difficulty experienced by surgeons working with a laparoscope is in understanding the conceptual relationship between:

(1) The camera angle and the port of entry of the laparoscope,

(2) The video view of the workspace as captured by the laparoscope, and

3

(3) The patient's anatomy.

Without this knowledge there is an increased probability of the surgeon losing orientation inside the workspace ultimately affecting patient safety (Crothers et al., 1999, Perkins N et al 2002, Dimitrios Stefanidis, 2006). Additionally the surgeon is also faced with issues of motion restrictions of the various tools and the limited field of view provided by the camera. As a result, handling the laparoscope to provide a view of the operating field is vastly more complicated than the simple, collocated relationship in open surgery, where the surgeon can look downward to directly see the simple, collocated relationship between the surgical tools and the patient. This difference also helps to explain why even surgeons who are experts in open surgery find it significantly more difficult to learn this conceptual relationship of MIS (Crothers et al., 1999). Proper camera orientation and navigation inside the workspace are important since they can impact the success/outcome of an MIS procedure (Kondorffer et al, 2004). In this research, we examine how these relationships are learned.

Current training systems have primarily focused on basic skill acquisition using rudimentary task setups. These are almost never duplicated in the OR. Secondly, current training techniques employ learning through rote memorization where the task is repeated until certain level of success is reached. While this type of learning is found to be helpful for teaching simple tasks, it has been found to be ineffective for teaching complex tasks (Bainbridge, 1997; Wulf & Shea, 2002) such as handling a laparoscope. In effect the novice surgeons find it increasingly difficult to perform just as well when required to adapt to the variations in the OR setup. This requires that trainer systems be designed such that effective transfer of learning is achieved. According to Van Merrienboier at al., (2006) the system must be designed such that it promotes the building of knowledge structures called schemas that are essential for abstractions and

4

elaborations of knowledge to find solutions to new problems. Also the training systems that facilitate the interpretive aspects of learning and actively avoid facilitating means-ends search by the learners for the tasks presented (Sweller at al., 1998) are found to be effective for studies that aim at transfer of learning. The laparoscopic task used in this research employ this training approach.

One of the primary objectives of this work is to train subjects to use a laparoscope in a way that would enable them to transfer their knowledge to new problems within the same or different operational setups. They should also be able to extend this knowledge to new tasks and environmental conditions. It is proposed that this training will be accomplished by exposing the learner to different camera positions and camera angle in the workspace during task performance. Task location and orientation within the workspace will also be varied. To facilitate this training, a custom built box-trainer system was constructed. The effectiveness of the training will be assessed using a testing task which will require the learners to perform the task in an environment that approximates the OR.

While insufficient learning in a MIS task can lead to poor performance, various studies have found that a major reason for performance degradation in the OR is that of stress (Linn; Bernard; Zeppa; Robert 1984). Surgical environments are characterized by very high task loads, and ambiguity and distractions that subjects surgeons to high levels of stress. Stress has also been found to play a significant role in the performance of surgeons, with undue levels of stress impairing judgment, decision making abilities and communication (Wetzel et al., 2006) resulting in performance degradation. In comparison to open surgery, the MIS environment is generally considered to be more stressful (Breuger et al, 2001) due to the increased complexity of the surgical tasks resulting from the increased difficulty of the tool handling component.

5

Performances in simulated MIS tasks of participants under stress have shown that stress can lead to impaired dexterity and increase in psychomotor errors (Moorthy et al., 2003) but MIS training curriculum do little to prepare them for the realistic and the highly stressful OR environments. Hence the second objective of this work is to examine possible methods of training that may improve the ability of surgeons to perform their psychomotor tasks under stress in the operating room. The stress environment will be created by inducing stressors artificially while the subjects perform their task.

To create an appropriately stressed environment in these experiments, the effectiveness of various environmental stressors was assessed. The stressors were physiologic in nature and included:

(1) Background noise and talking, and

(2) A requirement to work at arm's-length during task performance.

The effect of these stressors was assessed by measuring commonly used stress response variables; Heart Rate, Heart Rate Variability, and subjective stress responses (Köhn-Seyer G et al 1985, Ira GH et al 1963, Payne and Rick, 1986,Gunilla Krantz, Mikael Forsman and Ulf Lundberg 2004, Evans and Cohen 1987, Zeier, 1994). Each of these measures was validated in pilot work prior to use in the study. Equivalent stressors were then implemented in the subsequent training experiment. The impact of stress on performance was assessed as well as the effectiveness of biofeedback in managing that stress.

The main experiment of the thesis examines the ability of different training methods to prepare a learner for performing a camera navigation task under stresses similar to those expected in the operating room. The types of stressors were selected for the study so that they are not confounded with actually learning the task (like, e.g., task complexity). This will ensure the

effects of stress found in the study are clearer to interpret. The effectiveness of the methods was evaluated by assessing the performance on a testing task under a new stressor. The results have implications for training in a variety of situations involving hand-eye coordination and stress, not solely MIS.

Chapter 2

Development of an instrumented box trainer and a standardized camera navigation task

2.1 Introduction

MIS training during its nascent stages comprised of novices learning the skills by observing the expert surgeons perform in the OR .This mode of learning which was adopted from open surgical skill learning was widely accepted by the surgical community to be ill equipped to teach the skills specific to MIS and necessitated a need to develop better training models. The success of simulated training environments in the field of pilot training was seen as a possible option for surgical training .These systems with the use of computers allow learners to practice skills in virtual environments that approximate reality and have the capability to assess performance and provide feedback. But the high cost of technology and the lack of knowledge of the skills that are essential for being a successful MIS surgeon were the major reasons for hospitals not to implement these systems in their training curriculum. But with recent advances in computer

8

technologies, low cost training systems have been developed with capabilities to objectively assess performance and provide feedback to the learners. This has allowed the residents the opportunity to train and practice outside the operating room.

2.2 VR systems versus Box Trainers

Some of the approaches to training the fundamental skills of MIS include use of

1. Virtual Reality simulators (VR):Systems that contains 3D computer graphics

2.   Inanimate models: Human parts made from moldable materials

3. Video trainers (VT):Systems that have video cameras and monitors

Training on both box trainers (Fried et al., 1999; Scott et al., 2000) and virtual reality (VR) trainers (Seymour et al., 2002; Ahlberg et al., 2007) has been found to show improvement in the performance of junior surgeons in the operating room. In comparison tests between groups trained on either using a box trainer or a VR system, no significant advantage of one over the other in learning skills have been shown (E.C. Hamilton et al 2001, Munz et al, 2004, Jordan Newmark 2007). Unfortunately, there is not yet enough data on the effect of transference to the operating room for both together, let alone separately, to make a comparison between the two in that regard.

However, there are several practical advantages to developing a box trainer rather than a virtual reality system. These include

1. Cost: As the source code for virtual reality simulators are proprietary, there is a much longer time involved in developing a VR simulator versus a training box.

9

2. Lack of realism: Also, only relatively recently have many VR trainers been given haptic feedback, which is both very expensive and not particularly realistic. The latter is particularly important as friction characteristics vary largely among instruments and trocars (Dankelman, 2008).

3. Real time objective assessment: Until recently, the main disadvantage of box trainers is that they did not contain automatic performance assessment. However, the development of box trainers that can track laparoscopic instruments has alleviated this limitation (Datta et al., 2001, 2002; Dubrowski et al., 2005, 2007, Chmarra et al., 2006).

2.3 Review of camera navigation tasks in simulated training environments

Various studies have developed custom navigation tasks for teaching laparoscopic skills in both virtual reality and box trainer setups. In this section we will discuss the various tasks that have been designed to learn camera navigation.

2.3.1 Camera navigation task designed for VR systems

One of the first VR systems to teach visual spatial skills was the Endotower (Haluck et al 2001) which consisted of a 3D block tower with holes that held the targets. The learners were expected to find these targets by exploration using an angled laparoscope. Similarly (Eyal and Tendick 2001)in their study in virtual environment for training on angled laparoscope used long target boxes suspended in space and the task was to use the scopes to find specific letters inside these boxes.

In a study that aimed at evaluating a VR based system, LapSim, for training laparoscopic skills to novices, a camera navigation task was designed that required the subjects to use a 30 degree laparoscope to find and focus on a number of balls that appear randomly in a virtual environment (Hyltander et al 2002). A similar task was developed for evaluating the effectiveness of another VR system, LapMentor, where subjects were required to find and focus on 5 foam "dots" measuring 1 cm in diameter, which were placed at various preselected locations within the abdomen using an angled laparoscope (Pamela B. Andreatta 2006).

These tasks were designed to differentiate the novices with the experts. Objective assessment of performances such as errors, numbers of correct targets identified were the main determinants of level of proficiency.

2.3.2 Camera navigation task designed for box trainers

For teaching camera navigation tasks using box trainers Korndoffer and his colleagues (2005) developed a camera navigation task with a box trainer using a method similar to Eyal and Tendick (2001). Molinas and his colleagues (2008) developed the laparoscopic skills testing and training (LASTT) model to test camera navigation and other laparoscopic tasks using a box trainer. The camera navigation task requires the subject to identify targets which were mounted on the different modules in such a way that they could only be identified by moving the laparoscope in all directions (rotation, lateral and zoom-in/out movements). Each target included a large symbol only identifiable from a panoramic view and a small symbol only identifiable from a close-up view. In this study we will look at a camera navigation task that was designed for use with a box trainer.

2.4 Objective of navigation task design

One of the fundamental problems while performing MIS is in understanding the conceptual relationship that exists between the tools that need to be manipulated with the video view of the laparoscope and the patient's anatomy. Having a very good understanding of this relationship is essential for effective handling of the laparoscope. Current training systems have largely focused on skill acquisition by mainly employing simple task setups with no variations in camera angle or the port location even though these variables are known to significantly affect surgical performance. The main goal of these training systems is in acquisition of skills that are learnt by practicing the task on a specific setup and generalized to entire workspace rather than promotion of essential knowledge that could help the individual to handle variations in task setups. This type of training has found to be less effective when learners are required to perform the same task in different setups due to poor transfer of knowledge (G B Hanna et al 1997, Elspeth M. McDougall, MD 2006). Our training objective focuses on enabling the learners to develop a conceptual model of the complete hand-eye relationship, including the use of different positions of the camera and camera angle in the workspace.

For the actual task, we used target cylindrical boxes, in a similar manner to Eyal and Tendick (2001), suspended in space in different positions and orientations inside a box trainer and requiring the learner to use the different camera and port combinations to identify the targets in an effort to generalize the workspace.

12

2.5 Development and testing the trainer components for reliability

To achieve the training objective it is important that not only the training method but also the various components of the trainer system are designed and tested for quality. This section describes the development of the four main components of the training system:

1. The task box

2. The simulated laparoscope

3. The specific navigation task system

4. Measures of performance assessment

The main focus of this chapter concentrates on the development of two important components of the training system,

1. The task difficulty metric.

2. The collision detection system.

And the reliability tests performed as part of the standardization procedures.

2.6. Design of trainer components

2.6.1 Task Box Design

The main goal of the trainer design is that we want to be able to move the laparoscope easily between multiple port positions and the use of different lens angles. A box trainer (Figure 2.1), representing the human anatomical space, was developed for training and testing study participants. It consisted of a rectangular box with opaque sides to prevent study participants

13

from viewing the presented task inside the box .The top of the box was covered with an opaque cloth to simulate the opacity of the body that a laparoscope must be inserted through to reach the abdomen area. To simulate a port through which the laparoscope is inserted, a simulated trocar was constructed from plastic and a rubber sheet attached to velcro, 12mm in diameter and 20 mm in depth (Figure 2.2). To allow for placement of the port anywhere on the box, a slide arrangement (Figure 2.3) was created that could move the trocar to any desired position. A lock on the simulated trocar was created to fix it in space when training in that location was being performed. In our study, two locations center and left along the middle of the box was considered as the port of entry of the laparoscopes. Circular holes were cut in the cloth where the ports were placed.
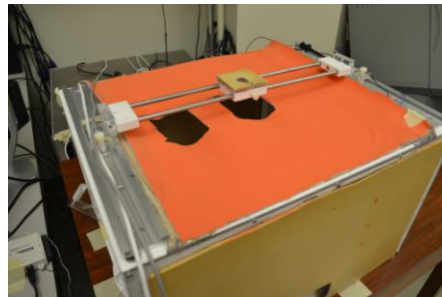
Figure 2.1:Box Trainer
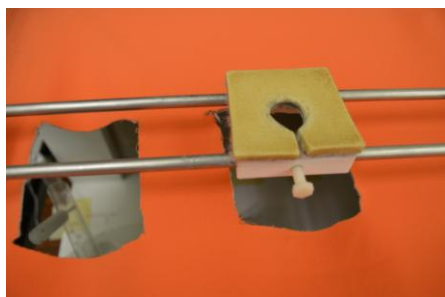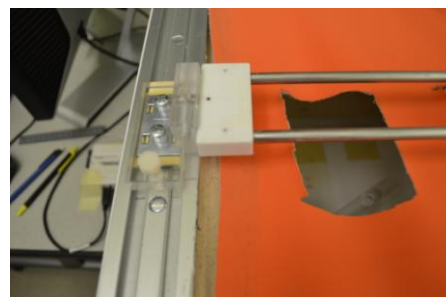
Figure 2.2: Port   Figure 2.3: Slide Arrangement

2.6.2 Simulated Laparoscope Design

A simulated laparoscope (Figure 2.4 and 2.5) was designed and built for the study which could be varied in its camera angle from 0 to 45 degrees. This was created by disassembling a 3DMed laparoscopic camera with a 0 degree camera angle. The camera was mounted on a pivot that was connected to a shaft 10mm in diameter and 35cm in length to simulate the laparoscope size. The pivot was connected to a knob at the handle of the simulated device to provide a way for a user to adjust the camera angle without having to disassemble the device.

Figure 2.4:0 degree Laparoscope   Figure 2.5: 30 degree Laparoscope

2.6.3 Navigation task Design

2.6.3.1 Physical Task Structure

The system used to train users in this navigation task was a tree with "branches". Each branch was a hollow cylinder, located at different positions on the tree and different orientations. Each cylinder contained a target at different depths to vary the difficulty of viewing inside one of the cylinders (Figure 2.6, 2.7, 2.8). The targets were rated as "Easy", "Medium", and "Hard" using a geometric metric to describe their required viewing position. This geometric metric is calculated based on the sum of the scores given for each of six parameters (Table 2.1).

15

Figure 2.6: Target Depth: Front   Figure 2.7: Target Depth: Middle

Figure 2.8: Target Depth: Rear

Table 2.1.Geometric metric parameters for defining the task difficulty

Geometric Metric Scores			
Parameters	Easy(1)	Medium(2)	Hard(3)
Depth (cm)	0-5	5-10	>10
Letter position	Front	Middle	Rear
Angle (deg)	0-5	5-25	25-45
Cylinder Location (cm)	0-10	10-20	>20
Obstacles	0-1	2-3	>3
Letter Proximity (cm)	<0.5	1-2	>2

2.7 Importance of Quantifying the Task tree

Training the MIS resident involves practicing simulations using mechanical models of the surgical environment. These models have advantages such as

1. Providing safe, realistic learning environments for repeated practice

2. Providing feedback and objective metrics of performance.

It is of utmost importance for designers to standardize the components of MIS training to ensure quality of the system. This would ensure that the performance assessed are reliable and valid .This section discusses one such component; a task difficulty metric that reliably quantifies a learner's true level of proficiency.

2.8 Testing the validity of a task difficulty metric developed for a camera navigation task

2.8.1Objective

The main objective of this study was to design a task difficulty metric and test the reliability of the metric.

2.8.2 Materials

The system used to train users in this navigation task was a tree with "branches". Each branch was a hollow cylinder, located at different positions on the tree and different orientations. It contained targets at different depths to vary the difficulty of viewing inside one of the cylinders. The targets were rated as:-

1. "Easy"

2. "Medium"

3. "Hard"

using a geometric metric to describe their required viewing position. The Geometric score for each category is the sum of the scores given for each of six parameters (Table 2.1). To validate these scores, these scores were compared to participants' subjective perception of difficulty. The subjective impression was rated on a visual analog scale (VAS). The VAS scale has a range of 0 to 12 divided into three divisions,

1. Easy: A VAS score range between of 0-4

2. Medium: A VAS score range of 4-8

3. Difficult: A VAS score of 8-12.

2.8.3 Methods

A total of nine subjects participated in this study. All the subjects had no prior knowledge and experience with using a laparoscope. All the subjects (7 males and 2 females) were right handed. Participants were briefed about the components of the system, the task tree, the target letters placed inside the cylinders, the simulated laparoscope and port locations, as well as the purpose of the task. A total of four task trees were used in this experiment. Each tree consisted of twenty targets for easy, medium and hard targets in total 60 targets that needed to be identified by each subject in each tree. The difficulty level for each target was defined by using the geometric metric parameters. The score for each target derived from the geometric metric can take a range from 0 to 12. The subjects were required to find the targets using the laparoscope on a task tree. Once the target was identified the subjects rated the difficulty level for finding the target on a

18

visual analog scale from 0 to 12. The final subjective score for each target is calculated by taking the average of all the individual scores of all the subjects using both the 0 and 30 degree.

2.8.4 Results

The percentage of targets found in each category is shown in Figure 2.9. Identification of the targets reduced with increasing target difficulty (96.75(Easy) vs. 78.75(Medium) vs. 57.5(Hard)). 92% with a (s: 2.1) of all the targets defined as easy using the geometric metric were rated as easy by the subjects, 82% (s: 6.1) of the targets defined as medium using the geometric metric were rated as medium by the subjects and 86% (s: 8.6) of the targets defined as hard using the geometric metric were rated as hard by the subjects. This indicates a high level of reliability of the task difficulty metric that was developed. The targets that were not found were not rated and not included in the final results. Though the unidentified targets were not included in the results the first result that targets were identified less with increasing difficulty validates the metric.
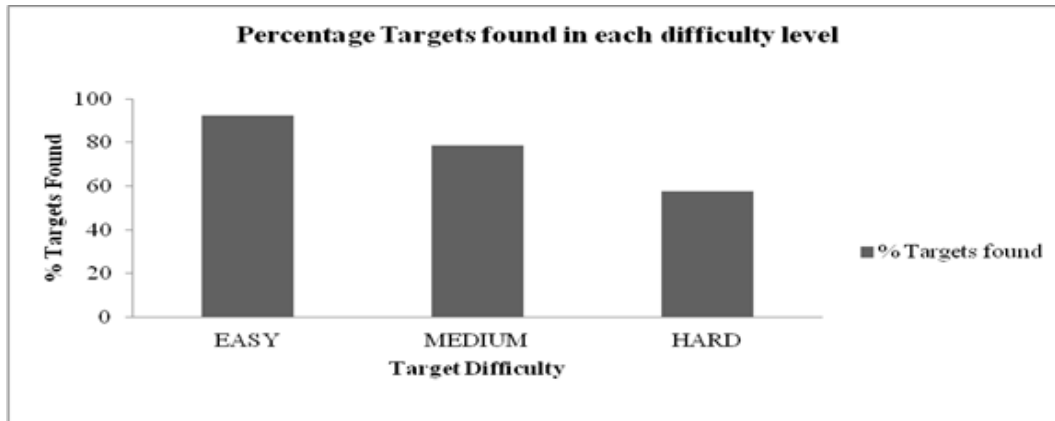
Figure 2.9: Percentage targets identified in each category

(Fig 2.10-2.12) shows the average VAS ratings for the easy, medium and the hard targets compared to the geometric metric. High correlation (r=0.89) was found between the geometric

19

metric and the VAS score indicating the reliability of this metric as a tool to control the difficulty of the task.
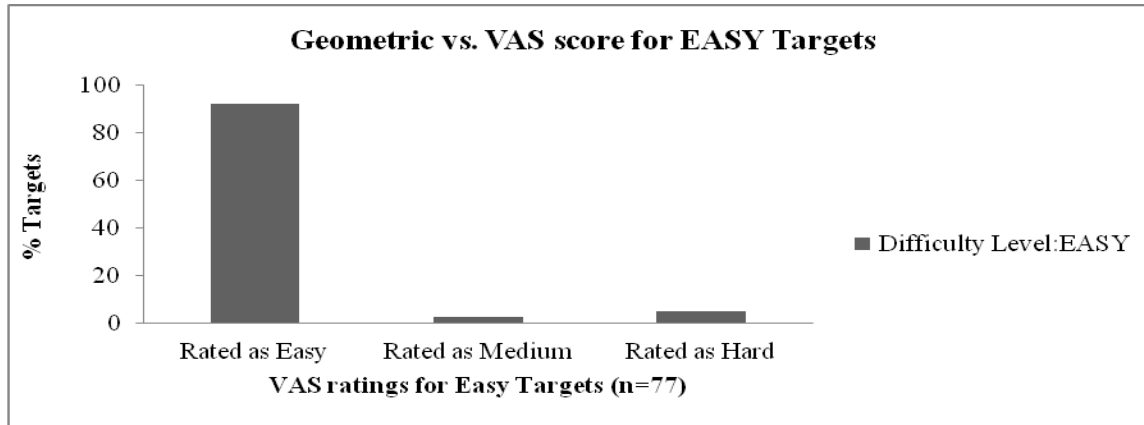
Figure 2.10: Distribution of VAS scores for the targets defined to be EASY by the Geometric metric
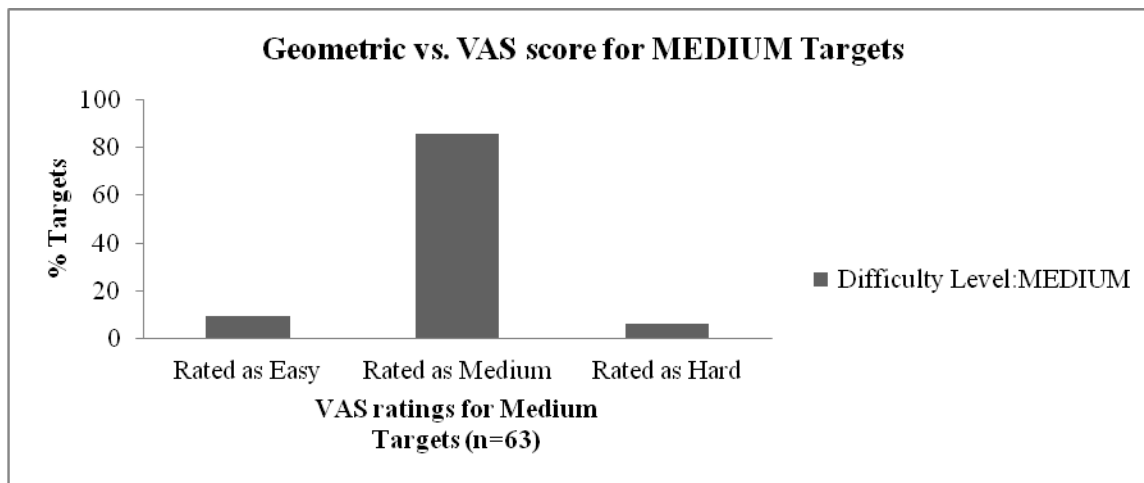
Figure 2.11: Distribution of VAS scores for the targets defined to be MEDIUM by the Geometric metric
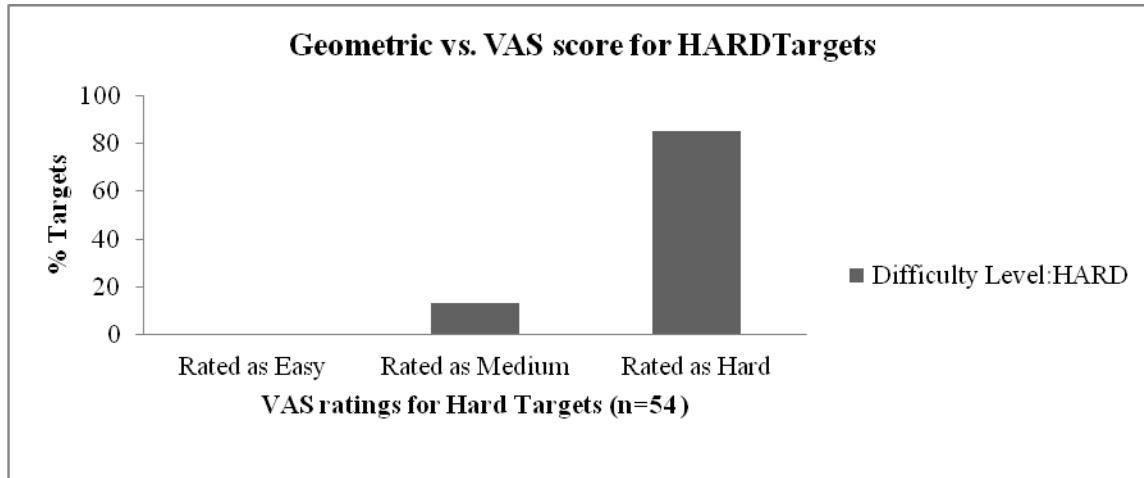
Figure 2.12: Distribution of target difficulty based onVAS scores for the targets defined as

HARD by the Geometric metric
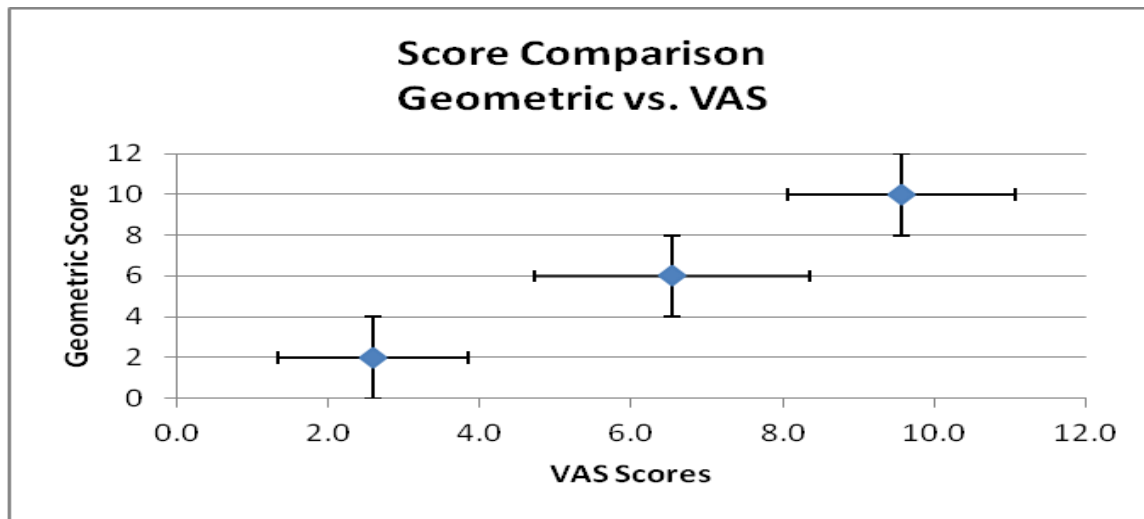
Figure 2.13: Geometric and VAS score comparison

2.8.5 Conclusions

Based on the results we were able to confirm that the geometric metric can be used as a reliable

tool for controlling the difficulty levels for the camera navigation task. Such a metric could

provide instructors with a reliable means to assess true performance levels as well as progress made by the subjects on a global scale.

2.9 Measures of Performance Assessment

Current performance measures in MIS training typically assess the learner's hand eye coordination by means of recording

- The economy of movement which is the ratio of the excess distance traveled to the optimal distance,

- Path length traversed by the instrument over time

- Smoothness of motion

- Number of errors

- Time to completion

Generally a comparative analysis is performed on the results of one or more of these measures at the end of training to determine the extent of learning.

2.9.1 Calculation of motion metrics

Currently the motion metrics are not fed back to the user while performing the task but are used for assessment purposes in terms of examining training performance. The advancement in computer has allowed for collecting real time data which allows us to process .In this study we recorded the various motion metrics using Motion Monitor. Motion monitor is an advanced data acquisition, analysis and visualization system developed by Innsport. The system consists of a sensor system flock of birds which are attached to the tool tip and a receiver which records the

22

position of the sensor in real world at 100 samples per second. The raw values are then processed by specialized software integrated with the system. The position data was then transformed mathematically to determine path length of the tool tip and smoothness of motion values. The total path length in 3D space was found by using the distance formula between two points in 3D space (x1, y1, z1) and (x2, y2, z2) and summing the individual distance of consecutive points over the entire duration of the trial.

D1 = √ ((x2-x1) ^2 + (y2-y1) ^2 + (z2-z1) ^2)

Total distance or Path length =∑D1, D2….D n where n is the last point recorded.

The second motion metric 3D Smoothness of motion was calculated by performing a third time-derivative of position in cm/s³.

$$j = \sqrt{(\frac{d^3x}{dt^3})^2 + (\frac{d^3y}{dt^3})^2 + (\frac{d^3z}{dt^3})^2}$$

Motion smoothness is derived from the integrated squared jerk as

$$J = \sqrt{\frac{1}{2} \int_0^T j^2 dt}.$$

Since for this study established motion metrics were used for assessment no specific tests were required to test their accuracy or reliablity.

23

2.9.2 Error Metric: Measurement of collision

2.9.2.1 Objective

One critical factor for a successful MIS surgery is to minimize unwanted contact of the laparoscope with neighboring tissues. Error as a measure of technical proficiency is a widely used performance metric in surgical training, as it is found to be reliable and valid in distinguishing the different skill levels of the learners.

The reliability of a metric is determined by the quality of the system measuring it. Hence it is a vital step in the development of surgical training system to ensure that the system is able to accurately measure a variable of interest. The collision detection system designed is an accelerometer that records changes in the g-force. The changes in g force can be used to record vibrations and this was the principle behind our design to use this property to record collisions on the task tree. Since sensors are prone to errors it is desirable to test the accuracy and reliability of the sensor to record changes which would ultimately greatly enhance the quality of the metric used. This section focuses on the test of reliability and validity conducted on the measuring system developed for recording collisions.

2.9.2.2 Materials

The system consists of a trainer box, task tree (a mechanical task component) and a collision detector consisting of a 3-axis accelerometer (Analog Devices ADXL 203).The accelerometer was placed inside the task trunk at half the height of the trunk as shown in (Figure 2.14). A rubber sheet (Rubbercal) was placed between the task tree and the bottom of the trainer box to act as a damper to prevent "knocking on the box" being recorded as a collision by the accelerometer. A Labview program was used for collecting the real time data. The collection

24

program assigned each collision as either major (>25gf) or minor (5-25gf) based on the intensity. It then displayed the results in real-time as the experiment was progressing.

Figure 2.14: Accelerometer inside the Task Trunk

2.9.2.3 Methods

The accelerometer was tested for its reliability to accurately record collisions. Intensity, frequency and location of collisions on different parts of the tree were considered as testing parameters.

2.9.2.3.1 Test 1.Testing the accuracy of the collision detector at different radial distances

The primary objective of this experiment was to test the reliability of the system to accurately record both minor and major collisions for various distances from the position of the detector. For each category of collision, that is to say minor and major, ten individual trials were carried out at each of the two distances 5 and 10cm from the detector. For each category of collision four repetitions were performed each with ten trials totaling forty. Accuracy for each distance was determined by dividing the total number of correct recordings by forty.

25

2.9.2.3.2 Test 2.Test of Reliability of the collision detector at various collision heights

The primary objective of this experiment was to test the reliability of the system to accurately record both minor and major collisions for various heights from the position of the detector. For each category of collision, that is to say minor and major, ten individual tests were carried out at specific heights (5cm, 10 cm and15cm) from the detector. There were in total of four repetitions for each height totaling forty trials. Accuracy for each height was determined by dividing the total number of correct recordings divided by forty.

2.9.2.3.3 Test 3.Test of Reliability of the collision detector at different tapping frequencies

The primary objective of this experiment was to test the reliability of the system to accurately record both minor and major collisions for various tapping frequencies. For testing the system for both minor and major collisions, tests were carried at specific tapping frequencies (1, 2, 3 and 4 taps per sec) at specific heights (5, 10 and 15 cm) and radial distances (5 and 10) comprised of ten individual taps. For a specific intensity (minor / major), a specific tapping frequency (1 2 3 or 4), a specific height (5, 10 or 15 cm) and a specific radial distance (5 or 10cm), four repetitions of ten trials were carried out totaling forty trials. Accuracy for each distance was determined by dividing the total number of correct recordings divided by forty.

2.9.2.4 Results

Figures (2.15-2.20) shows a plot of accuracy of the collision detection system for the various tapping frequencies, radial distances and heights.

2.9.2.4.1 Accuracy based on radial distances

The accelerometer system was able to record the minor and major collisions with accuracy greater than 90% for the different radial distances tested. The accuracy of the system to record minor collision was greater at 5 cm compared to 10cm from the sensor. The accuracy of the system to record major collision was higher for 10 cm compared to 5 cm from the sensor. Overall we determined that the accuracy of the system to record major collisions was higher when compared to minor collisions (96 (major) vs. 91(minor)) for the various radial distances

2.9.2.4.2 Accuracy based on heights

The accelerometer system was able to record the minor and major collisions with an accuracy of greater than 90% for the different heights tested. It was observed that for both minor and major collisions the accuracy rate was lowest for the height 15cm. Overall we determined that the accuracy of the system to record major collisions was higher when compared to minor collisions (96 (major) vs. 94 (minor)) for the different collision heights.

2.9.2.4.3 Accuracy based on tapping frequencies

The accelerometer system was able to record the minor collisions with accuracy of greater than 90% and for major collision with accuracy greater than 95% for the different heights tested. The accuracy of the system to record minor and major collisions was highest for tapping frequency 1 tap/sec and lowest for 4 tap/sec. Overall we determined that the accuracy of the system to record major collisions was higher when compared to minor collisions (98 (major) vs. 97 (minor)) for the different tapping frequencies. The difference was not found to be statistically significant. The damper material prevented recording collisions on objects other than the task tree.

27
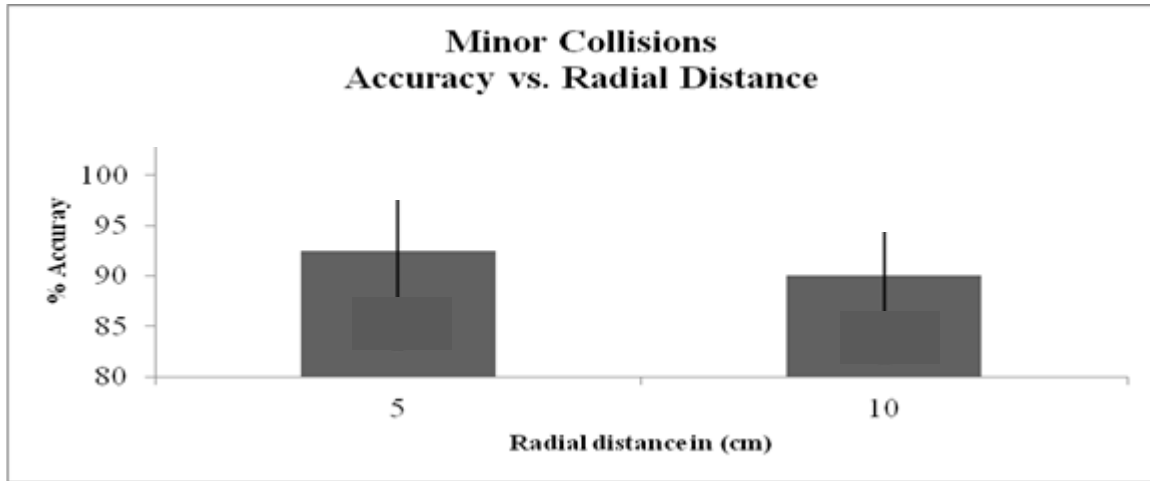
Figure 2.15: Collision detection accuracy at different radial distances for minor collisions
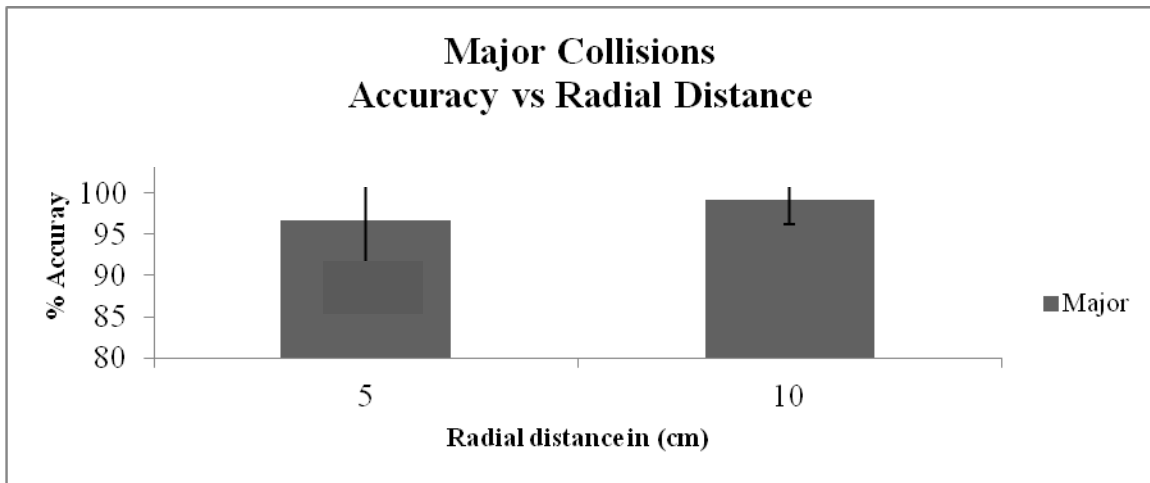
Figure 2.16: Collision detection accuracy at different radial distances for major collisions
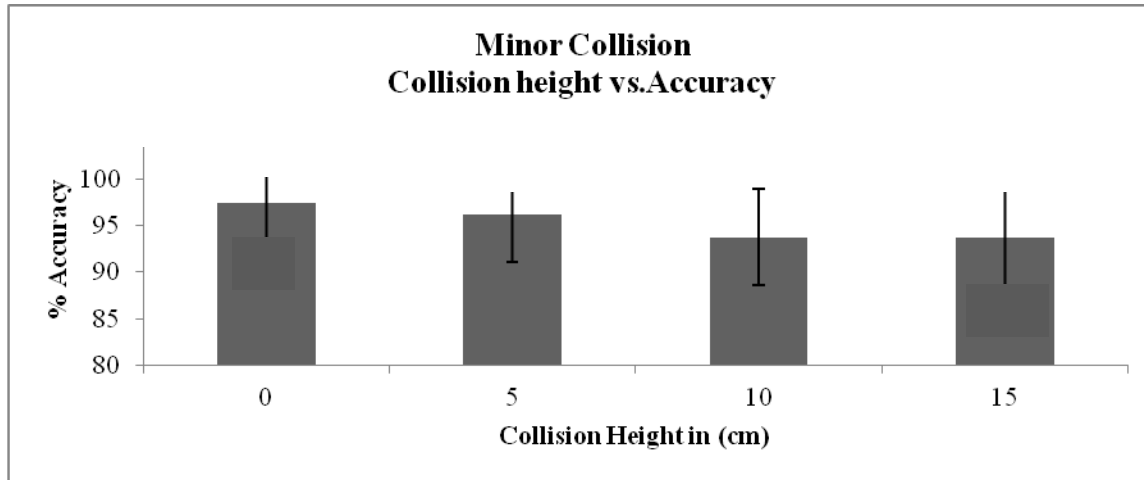
Figure 2.17: Collision detection accuracy at different collision heights for minor collisions
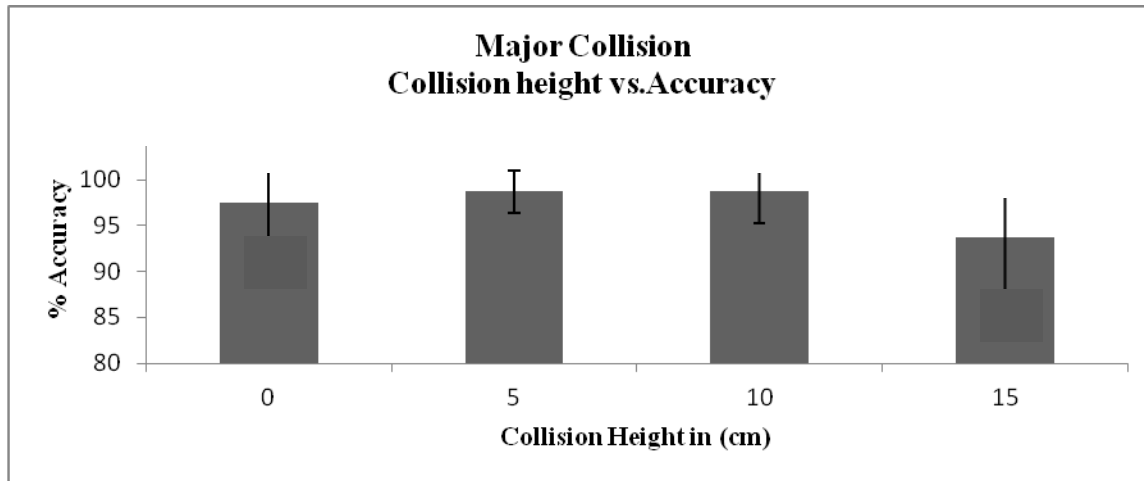
Figure 2.18: Collision detection accuracy at different collision heights for major collisions
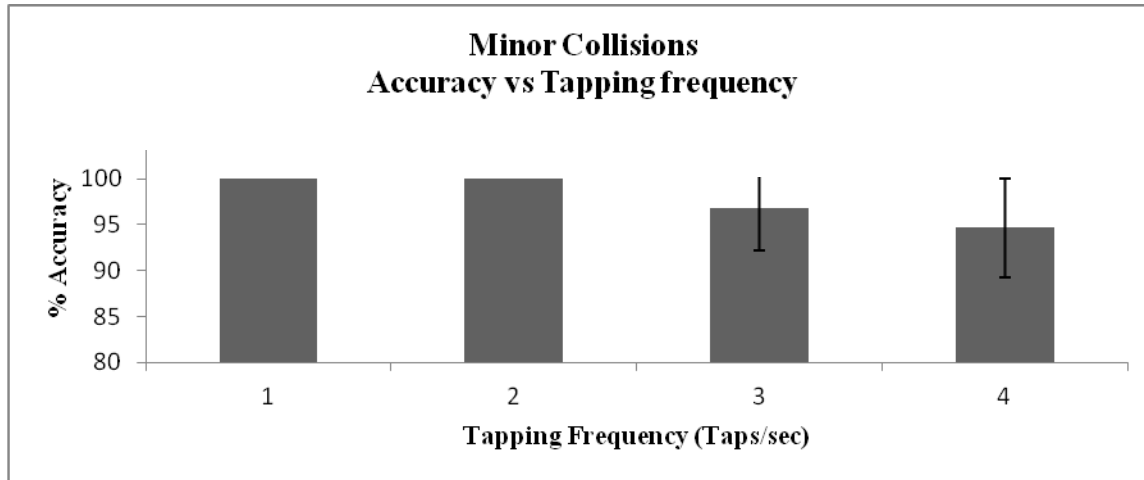
Figure 2.19: Collision detection accuracy at different tapping frequency for minor collisions
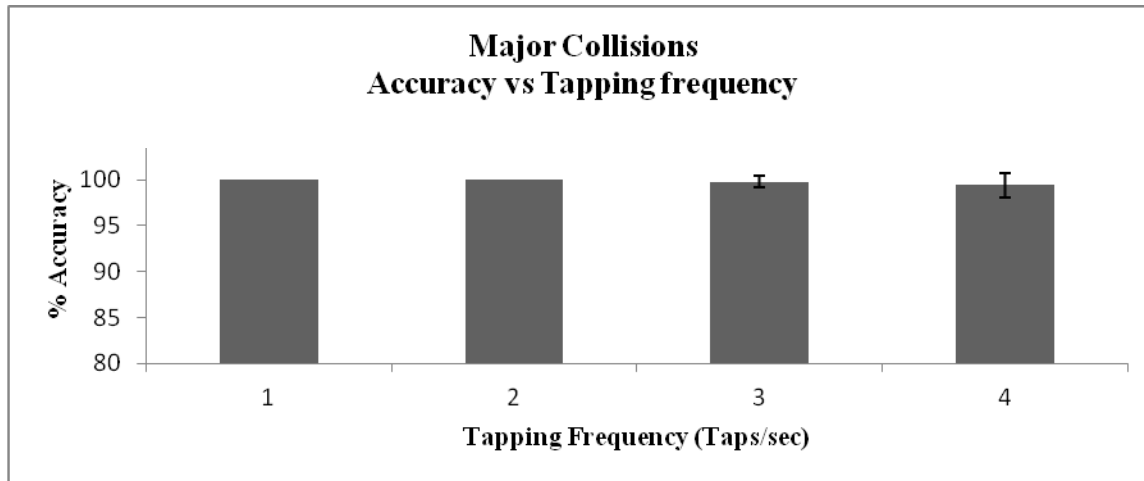
Figure 2.20: Collision detection accuracy at different tapping frequency for major collisions

2.10 Conclusions

The real time accelerometer based collision detection system has been found to be very accurate and reliable for collision detection measurements, classified into major or minor, in MIS training regardless of where in the mechanical task set-up a collision is made.

المنارة للاستشارات

www.manaraa.com

Chapter 3

Assessing techniques for measuring stress levels during a MIS task

3.1 Introduction

Stress refers to the reaction of an organism to a perceived threat (Seyle, 1930). It is an appraisal process in which perceived demands exceeds an individual's physical, emotional and cognitive resources (Driskell & Salas, 1996) resulting in several undesirable effects such as:

1. Physiological: Increased heart rate ,labored breathing, trembling (Rachman,1983)

2. Psychological: Fear, anxiety and frustration(Driskell and Salas,1991b),Motivational losses(Innes and Allnutt,1967)

3. Social and behavioral outcomes : loss of team perspective (Driskell, Salas and Johnston 1995), decrease in pro-social behaviors such as helping (Mathews and Canon,1975)

31

Evidence indicates that stress is a costly health-related issue, in terms of individual performance and well being, as well as organizational productivity (Ilgen, 1990). In order to mitigate the risk of stress on individuals, training to handle stress has become mandatory in professions such as in the military and aviation where stress levels are generally very high.

Stress is also an important issue to address in the field of surgery, particularly for MIS, which has higher levels of stress than open surgery (Berguer, 2001). To handle stress in MIS, the idea motivating the work in this section is to obtain quantitatively comparable stressors for use in further training and testing. In addition, if reliable, real-time stress measurements can be made, we could expose learners to a stressful environment while performing a MIS task and then to use biofeedback to help develop coping skills. For this, two aspects of this training are considered. First, one needs to be able to effectively induce stress, and a variety of stressors will be examined and compared. Second, to provide biofeedback, one needs to be able to continuously monitor stress levels. Heart rate (HR) and heart rate variability (HRV) are two potential biofeedback parameters which we examined to determine their accuracy in reflecting subjective stress.

3.2 Measuring stress

3.2.1 Physiological measure of stress

Stress triggers changes in the body releasing hormones to respond to the cause. When the body is exposed to a stressful condition, the autonomic nervous system releases epinephrine and norepinephrine producing what is known as the flight or fight response. The hypothalamic-pituitary-adrenal axis (HPA), part of the neuroendocrine system, releases corticotrophin hormone (CRH) activating the hypothalamus. The hypothalamus then secretes adrenocorticotrophic releasing hormone (ARH). This hormone stimulates the pituitary gland which is situated below

32

the hypothalamus to secrete Adrenocorticotrophic hormone(ACTH).This hormone in turn stimulates the adrenal or the suprarenal glands situated on the top of the kidney to secrete the stress hormones, namely adrenaline and urinary cortisol. These hormones produce certain functional adjustments such as heightened cardiovascular response to supply more blood quickly, diversion of the blood from less vital to more vital organs, breakdown of glycogen stores in the liver and muscle to get more glucose, and formation of more glucose from non carbohydrate substances helping the system to control the source of stress. The specific end effects of these bodily reactions are increased systolic and diastolic blood pressure, heart rate, urinary epinephrine and norepinephrine, salivary cortisol, as well as trapezius muscle activity, measured by surface electromyography. All of these have been found to be good indicators of stress (Gunilla Krantz, Mikael Forsman and Ulf Lundberg 2004). Thus, measuring these changes could prove useful for objectively measuring stress. We will consider two easily measured effects: heart rate (HR) and heart rate variability (HRV).

3.2.1.1 Heart rate

Heart rate refers to the number of heart beats per unit time. It is measured using an electro cardiograph which measures the number of R-R wave intervals per unit time. Heart rate is under the control of the autonomic nervous system (ANS) and follows changes of the ANS. The ANS has two main branches, the sympathetic and the parasympathetic nervous system. The sympathetic nervous system is involved in the control of the internal organs and helps the body to adapt to the external environments such as stress or exercise. The parasympathetic nervous system on the other hand works to bring the body to rest. Thus the SNS and PNS counteract each other to maintain homeostasis. One of the major factors that can offset this balance is stress, both mental and physical.

33

3.2.1.1.1 Effects of stress on Heart rate (HR)

Under stress there is an increased activity of the sympathetic system and decrease in activity of the parasympathetic due to the release of epinephrine and norepinephrine by the autonomic nervous system. This results in an increase in the heart rate from its resting state.

Numerous studies have shown that the heart rates of surgeons significantly increase during surgery or other stressful tasks (Foster et al 1978, Woitowitz et al 1972, Köhn-Seyer G et al 1985, Ira GH et al 1963, Payne and Rick, 1986). In other work environments, various groups have studied the impact of stress on heart rate and blood pressure (Dobkin and Pihl, 1992, LForsman and Lindblad, 1983, Taelman et al, 2009) and found that the mean heart rate to be higher for subjects under the influence of mental stress as compared to the non stress group. Heart rate has also been used to monitor stress in studies trying to understand the effect of workplace stressors such as noise and ergonomic constraints on performance (Lusk et al 2004). Other studies in which mathematical models have been developed for predicting heart rate and oxygen consumption under varying thermal load and workloads have also found heart rate to positively correlate with stress (Mercenkaya.R 1974).

Heart rate thus can be a good indicator of stress; however, it is subject to other factors like fatigue, posture, caffeine and sleep. Thus the results can be confounded by factors other than the primary stressor. Therefore, when using heart rate as a measure, consideration must be given to the impact of extraneous variables like body position and environmental conditions that can potentially distort the readings.

3.2.1.2 Heart rate variability

Heart rate variability (HRV) is another variable that is used to measure the stress on an individual. It is a measure of the variation in the R-R intervals over time. The control center in the medulla oblongata called the nucleus ambiguous controls the parasympathetic activity of the heart through the vagus nerve during the respiratory cycle. The vagus nerve is excited during expiration and inhibited during inspiration which in turn inhibits the vagal nerve stimulation of the SA node which is responsible for the normal sinus rhythm. Thus a periodic variation is exhibited in the beat to beat intervals resulting from the breathing cycle. This phenomenon, referred to as Respiratory sinus arrhythmia (RSA), is found to change under stress.

3.2.1.2.1 Effect of Stress on Heart rate variability (HRV)

Stress has been found to reduce the vagal tone, which eliminates the variability that was exhibited under normal conditions. This physiological phenomenon has found use in the objective measurement of stress. HRV, as a measure of mental stress, has been used in various studies (Hjortskov.N et al, 2004). Studies trying to identify the effects of stress on surgeons found a positive correlation between HRV and work stress (van Amersfoort, 1999, Langelotz et al, 2008). HRV was also used as a measure of stress to address the high stress levels encountered by surgeons while performing laparoscopic surgery as compared to conventional surgery (Bohm et al 2001).

3.2.1.3Methods of measuring HRV

3.2.1.3.1Time domain

HRV is can be calculated in either the time domain or the frequency domain. In the time domain there exists an inverse relationship between HRV and stress, that is, the higher the stress, the lower is the variability. Time domain calculation of HRV to measure stress calculate the inter-beat variability (time between two R-R of the QRS waveform) and adjacent cycle lengths using

1. Standard deviation of NN interval (SDNN): The standard deviation of NN intervals calculated over the entire duration of the trial

2. Square root of the mean squared differences of successive NN interval (RMSSD) : The square root of the mean squared difference of successive NNs

These measures of HRV (SDNN and RMSSD) have been found to be good indicators to measure the effect of autonomic tone on the heart due to physical and psychological stressors (Kleiger RE, 1992).In one study it was found that SDNN was able to clearly differentiate the different levels of stress (MG Kang et al 2004).

In this study to measure stress using HRV the time domain variables SDNN and RMSSD will be calculated using the Kubios HR system to determine if they can be used as a reliable biofeedback variable.

3.2.1.3.2 Frequency domain

In the frequency domain analysis of HRV, the power spectral density (PSD) provides basic information on the power distribution across frequencies. Two main frequency components are used to study the effect of stress

36

1. Low frequency component: Power (ms^2) in the frequency range of 0.05-0.149 Hz

2. High frequency component (Respiratory Sinus arrhythmia (RSA)): Power (ms^2) in the frequency range of 0.15-0.40 Hz.

The ratio between LF/HF is normally used to determine the physiological effect of stress.

Low-frequency HRV is mainly used to index sympathetic nervous system activity and high-frequency HRV is related to parasympathetic (vagal) activity. These results stem from the fact that low frequency HRV are found to be abolished by sympathetic antagonists and high frequency HRV are abolished by parasympathetic antagonists. It is observed that under stress there is a reduction in the high frequency components due to the reduction of the parasympathetic activity and increase in the low frequency components due to the sympathetic activity.

Studies have revealed that the low frequency component of HRV to positively correlate with increased mental stress (Wang et al 2008). Other findings of significant correlations between depressions, anxiety, and emotional stress to low HRV validate its usage as a good estimator of stress (Dishman, et al., 2000, Carney, et al., 1995). In this study we both the LF and HF components will be calculated and will be used to determine if high correlation exists between these values and stress.

Heart rate measurements can be a useful measure of stress if the changes in the heart rate are indeed a result of undesirable psychological or physiological state. However, HRV as a stress measure is influenced by different factors. The time domain measure of HRV is greatly influenced by respiration and is found to increase with increased respiration rate and increased depth of respiration (Atsuo Murata 1992).Other factors such as posture also affects heart rate

37

variability in that HRV was found to be higher for standing and walking tasks compared to sitting tasks. These findings have provided researchers with the necessary insight into the variables one needs to be cautious about when performing new experiments (I.Sipinkova 1997).

3.2.2 Subjective Measures

The basis of self report methods of stress measurement is based on the theory that stress affects how we perform, how we feel, and many of our bodily functions. According to the psychological model of stress, humans use their cognition to perceive stress and appraise the situation. Subjective methods have been employed for the measurement of stress and mental workload as an alternative to physiological measures since physiological measures are prone to external factors other than the stress inducing factors.

Studies have found that self report measures are preferable in the case of cognitive and emotional stress assessment due to the greater face validity and reliability (Zeier, 1994). The other advantages of subjective methods are that they are easy to administer and are inexpensive. In stress measurement studies, they have been found to be better predictors of stress (Evans and Cohen 1987) which allows this non invasive method to be used in this study. However, subjective assessments have certain drawbacks. The reliability of these methods can be affected by the individual responses which could be corrupted by cognitive inconsistencies and personality traits. However self reports methods have been found to be effective and valid if the questionnaire is well designed (Frese & Zapf, 1988, Ganster, Dwyer & Fox, 1993, and Belkic et al ,2007). In this study we used one such subjective measurement tool called a visual analog scale for subjective assessment of stress during the task. This measurement technique has been found to be easier to administer, has a better responsiveness and also a highly reliable and valid scale (McCormack HM et al 1988, Bond A 1974, Aitken RC et al 1969).

3.3 Determination of effective stressors for a laparoscopic camera navigation task

3.3.1 Objective

The primary purpose of this study was to determine stressors that could produce the same level of stress for use in the main study. The idea motivating the work discussed in this section is to obtain quantitatively comparable stressors for use in further training and testing. The study was also designed to determine the reliability and validity of the various physiological stress measurement variables to measure stress. In addition, if reliable, real-time stress measurements can be made, we could expose learners to a stressful environment while they are performing a MIS task and then to use biofeedback to help develop coping skills. To achieve this, the following two aspects of training are considered.

- Create a simulated stressful environment: To achieve this, a variety of stressors that are commonly experienced in the OR will be examined and compared.

- Use of biofeedback to continuously monitor stress levels: To accomplish the second objective we examined two potential biofeedback parameters:-

    1. Heart rate (HR): Frequency of repetition of the P-Q-R-S waveform over time. Measured in (beats/minute)

    2. Heart rate variability (HRV)

These two variables will be tested for their accuracy in reflecting a subject's perceived stress levels.

3.3.1 Materials

The box trainer with the camera navigation task was used for this experiment. Heart rate was recorded using a Noraxon Myosystem and heart rate variability (HRV) was calculated using Kubios HRV Analysis Software.

3.3.2 Methods

A total of nine subjects participated in this study. All the subjects had no prior knowledge and experience with using a laparoscope. All the subjects (eight males and one female) were right handed. A total of five task trees were used for the study; each task tree consisted of a total of twelve targets. The task involved identifying as many targets/letters as possible located at different depths and orientation on the tree (in 3-D space) using the simulated laparoscope, while minimizing the number of collisions with the task structure. The task was performed under four potential stressors condition (Figure 3.1-3.5) and a no-stress condition.

1. **Arm-stretch:** It involved the non dominant hand of the subjects to be raised at shoulder height for the duration of the trial. The arm stretch stressor is a kinematic stressor.

2. **NMES:** Neuromuscular electrical stimulator (NMES): This stressor creates pain and discomfort by exciting the muscles of the non-dominant hand of subjects by passing small amount of current (varying between 8-40mA) through it. The pulsating current will relax and contract the muscles periodically creating discomfort and inducing stress. This is a physiologic stressor.

3. **Balance:** Subjects were made to stand on an unstable platform for the entire duration of the trial. They were asked to maintain their balance and were not permitted to rest. The balance stressor is a kinematic stressor.

4. **Noise:** Participants were subjected to a summation of continuous 90db pink and white noise over the entire duration of the trial. This noise stressor is a physiologic stressor.

5. **No stress (NS: control)**

Stress measurement variables, heart rate and heart rate variability, were recorded during the entire duration of the task. The subjects were also required to rate their stress levels on an analog scale based on how they perceived it. Finally a comparative analysis was done between the recorded variables for the various stressors to determine the most effective stressor and the best stress measurement method. The task trees and the stressors were randomly presented to a subject, being counterbalanced across subjects. Heart rate and heart rate variability were recorded for the entire duration of the task with each stressor. The subjective assessment of stress was recorded on an analog scale from 0 to 10 after the completion of each tree. HR and HRV were compared to the subjective stress to assess reliability.
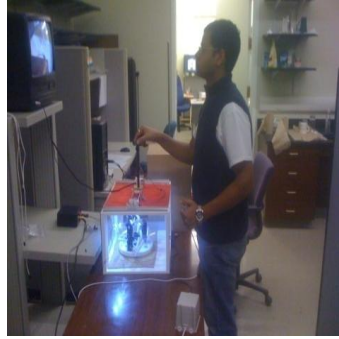
Figure 3.1 No stress　　　　　　　Figure 3.2 Arm stretch

Figure 3.3 Balance　　　　　　　Figure 3.4 Noise

Figure 3.5 NMES

3.4 Results

3.4.1 Subjective stress index and stressors

The results from the study indicated that subjects perceived the stress levels induced by Arm
stretch, Noise, NMES and Balance stressors to be higher than that of the No stress conditions.
(Arm: 6.21±1.54, Balance: 6.93±1.13, NMES: 6.17±1.01, Noise: 4.82±1.97 vs.NS:3.28±2.24) as
shown below.
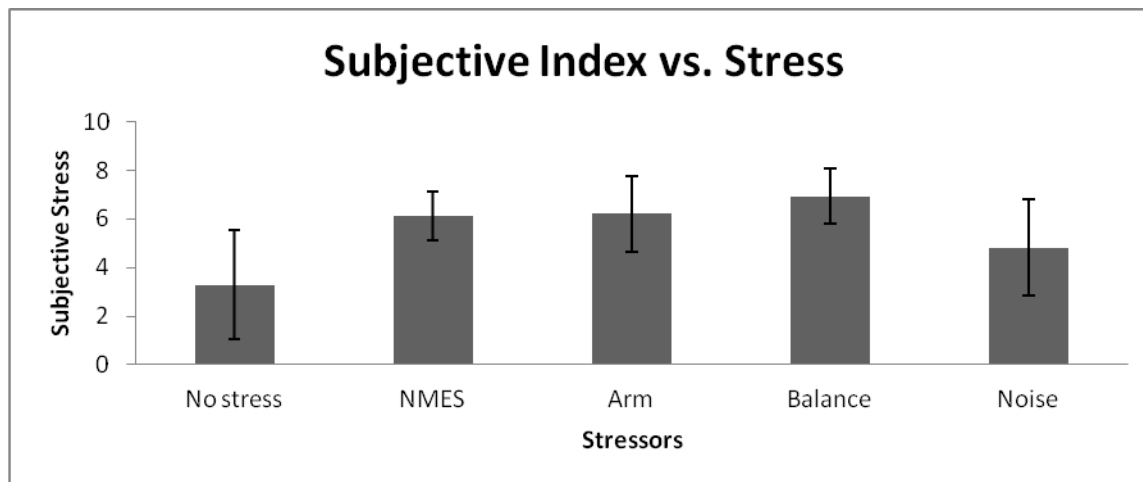
Figure 3.6:Comparison of Subjective stress index against the stressors

3.4.2 Stress and HR

The mean heart-rate was found to be higher for the stress conditions as compared to the non
stress condition(S: 98±13 bpm vs. NS: 92±11 bpm) .The graph below shows the average HR for
the various stressors.
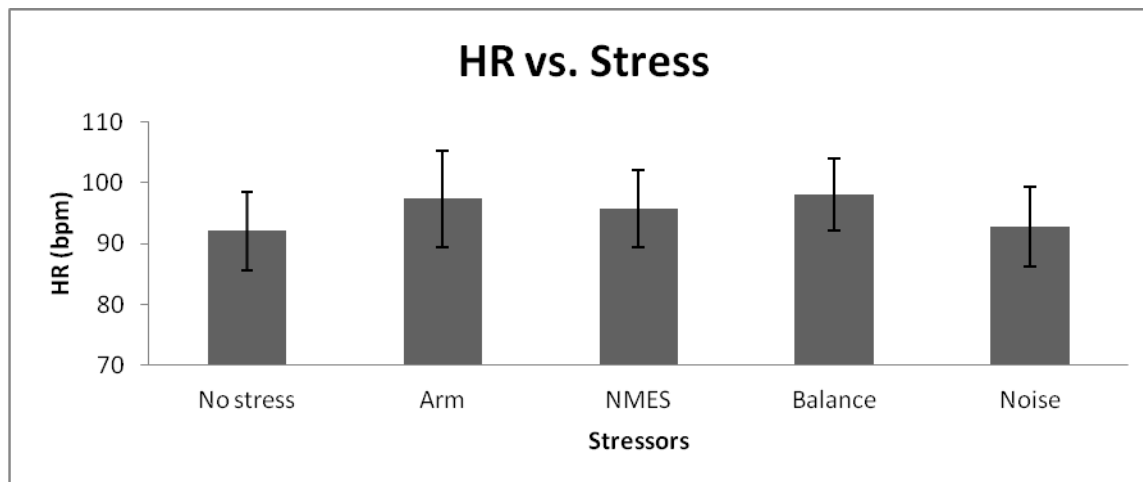
Figure 3.7:Comparison of heart rate against the different stressors

Correlation between subjective stress index and average HR for the different stressors are as below

- Positive and low for the Arm stretch(r= 0.10)

- Positive for NMES stressor(r= 0.23)

- Negative for the Balance (r= - 0.47)

- Negative and low Noise stressor(r= - 0.09).

- Positive and low for No stress (r= 0.22)

3.4.3 Stress and HRV

3.4.3.1 Time domain measure: RMSSD

The time domain measure of heart-rate variability (RMSSD) was found to be less for stress compared to the non stress conditions(S: 25 ± 13 ms vs. NS: 28 ± 23 ms). The graph below shows the average RMSSD for the various stressors.

44
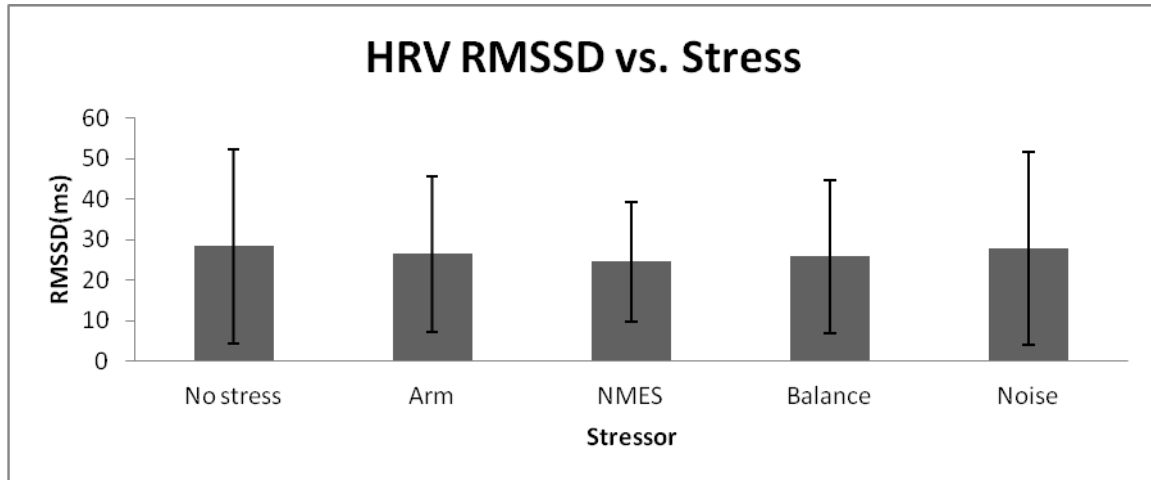
Figure 3.8: Comparison of HRV. (RMSSD) measures against stressors

Correlation between subjective stress index and time domain HRV (RMSSD) measure for the different stressors are as below

- Negative and low for the Arm stretch(r= - 0.12)

- Negative and low for NMES stressor(r= - 0.08)

- Negative and low for the Balance stressor (r= - 0.049)

- Negative and low Noise stressor(r= - 0.06)

- Negative and low for No stress(r= - 0.14)

3.4.3.2 Time domain measure: SDNN

The time domain measure of heart-rate variability (SDNN) was found to be slightly less for stress compared to the non stress conditions(S: 50.08 ± 24 ms vs. NS: 50.62 ± 23 ms). The graph below shows the average SDNN for the various stressors.

45
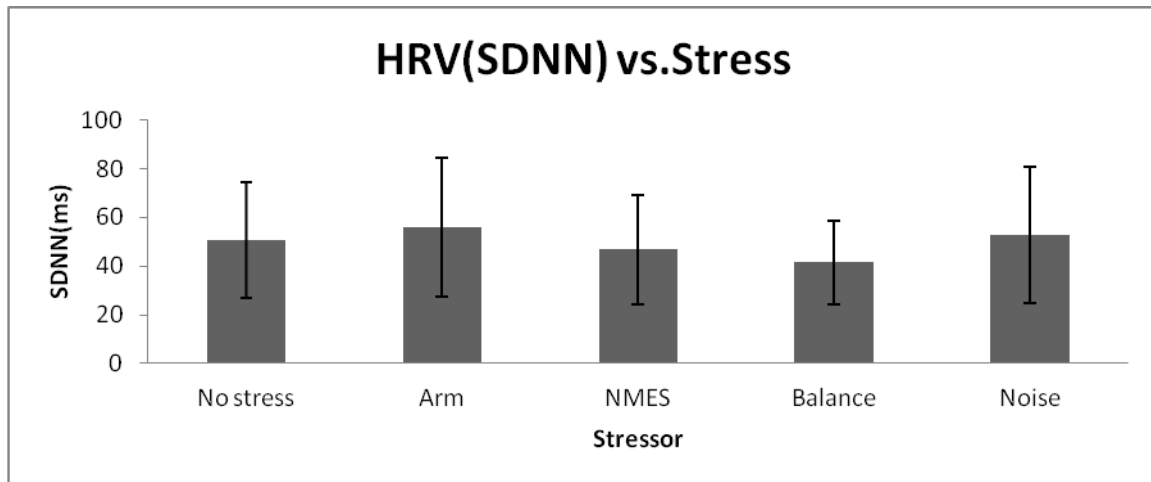
Figure 3.9: Comparison of HRV. (SDNN) measures against stressors

Correlation between subjective stress index and time domain HRV (SDNN) measure for the different stressors are as below:

- Positive and low for the Arm stretch(r= 0.13)

- Positive and low for NMES stressor(r= 0.16)

- Positive and low for the Balance stressor (r= 0.18)

- Positive and low Noise stressor(r= 0.135)

- Negative and low for No stress(r= - 0.20)

3.4.3.3 Frequency domain measure: LF/HF

The frequency domain measure of heart-rate variability (LF/HF) was found to be slightly higher for stress compared to the non stress conditions(S: 5.26 ± 2.94 vs. NS: 5.07 ± 4.44). The graph below shows the average LF/HF for the various stressors.

46
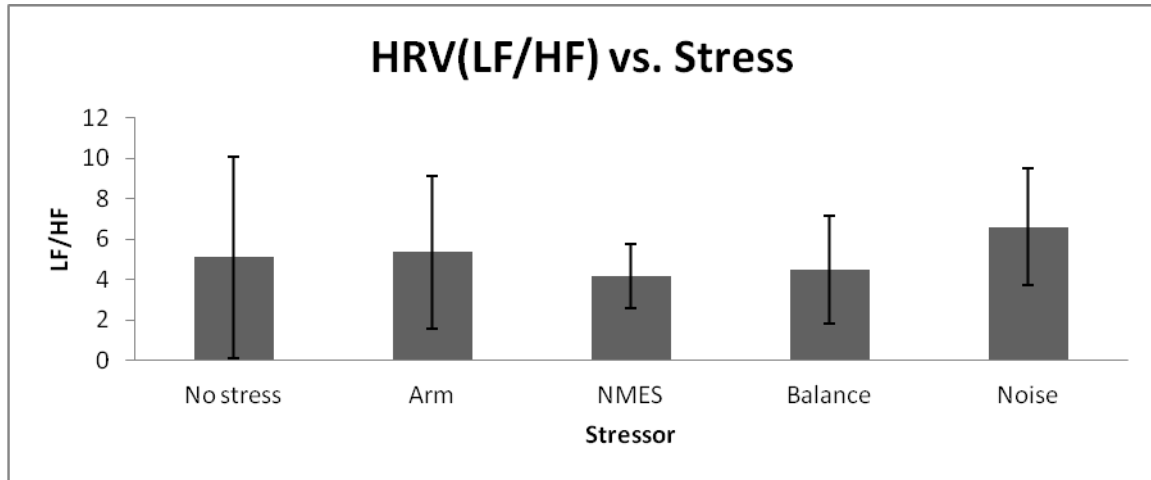
Figure 3.10: Comparison of HRV (LF/HF) measures against stressors

Correlation between subjective stress index and frequency domain HRV (LF/HF) measure for the different stressors are as below:

- Positive and low for the Arm stretch($r= 0.13$)

- Positive and low for NMES stressor ($r= 0.27$)

- Negative and low for the balance ($r=- 0.16$)

- Positive for Noise stressor($r= 0.50$)

- Negative and low for No stress ($r= - 0.017$)

While there existed a clear difference in the subjective stress index scores for the various stressors, correlation performed between subjective and objective measures (HR and HRV)of stress for the various stressors was inconclusive.

47

3.5 Conclusions

Even though the mean HR and HRV showed the correct trends when comparing the stress conditions to the no stress conditions (i.e., HR increased and HRV decreased), there was a poor correlation of HR and HRV to subjective stress index. The results indicate that the physiological measurements are not sensitive enough to record perceived stress for the particular conditions of the experiment. From these results, we have found that HR and HRV would provide a poor indication of stress and would not be effectively used as biofeedback parameters. The average HR and HRV to record stress during the trial duration could have been affected by periods of high and lower stress that could have had a reduced the overall scores for these variables. We did however observe the HR and HRV varying over the entire duration of the trial.

The different stressors were also assessed in terms of their ability to produce stressful conditions based on the subject stress evaluations. This is particularly necessary as in Chapter 4, we will present an experiment that requires the controlled manipulation of stress conditions both in training and testing.   From the current experiment, the participants perceived the no stress condition as being relatively non-stressful. The nonzero value was not surprising as the base task was somewhat stressful in itself. For the noise stressor, participants initially found it to be more stressful than the base task, but this value lowered over time resulting in a score comparable to the no-stress condition. This is likely due to the fact that humans have been found to adapt to noise over time (Loewen 1992). In contrast, the subjects rated the balance stressor to be the most stressful. This is presumably due to the unstable condition in which they were subjected to perform the task. However, the balance stressor also resulted in the poorest performance as determined by the number of targets identified (4 vs. 7) and the number of collisions (16 vs. 5.5) compared to all other groups. This suggests that a too high a stress level had been induced:

48

researchers have found that very high stress levels can interfere with learning by overloading memory and preventing learning of basic skills (Keinan and Friedland, 1992). Thus, we considered this stress level too high to be used for stress induced training. The stress scores for the arm stretch and NMES were identical. They were likely found to be stressful for the subjects as the stress construct was the pain and discomfort that each stressor induced. Based on the results of our study we identified the arm stretch and NMES stressor to be used as the stressors for the study .The stress due to these are high enough but did not adversely affect performance indicated by the targets identified and collision count and would allow us to simulate the stressful conditions the surgeons experience in the OR.

Chapter 4

Evaluating the Effect of Different Training Methods on Handling Stress during a MIS Task

4.1 Introduction

Stress is a major factor for performance deterioration in surgical environments, particularly minimally invasive surgery (K. Moorthy 2003,Wetzel C et al 2006, and Arora et al 2010).   In general, stress affects an individual's physiological, cognitive and emotional processes which in turn have a direct impact on task performance. These observations suggested that it is important to train individuals to overcome the negative effects of stress. The drawback of current training systems for minimally invasive surgery is that they do not consider actively training for stress in their curriculum even though it has been found to degrade performance. Being a profession where even the slightest mistake could be potentially fatal, it would be of

great benefit to effectively train surgeons who would be able to manage stress effectively, and thereby avoid undesirable performance deterioration. The main effort of this work was to evaluate the impact of three different training methods on handling a laparoscopic task under stress.

Current MIS training systems typically only use performance measures to assess the proficiency levels of the learners. We will consider the effectiveness of this standard training method to two others that could be potentially more effective in handling stress. The first of these methods is training not only for high levels on the performance measures but also for low cognitive load. The premise of this method is that training for a low cognitive load is an indication of the automaticity of the task. This training method has been found to be effective for maintaining effective task performance in stressful environments (Driskell, Willis, & Copper, 1992, Cascio, 1991, Deese, 1952, Fitts, 1965; Weitz, 1966). Since stress can affect the cognition of an individual (Finkelman, Zeitlin, Romoff, Friend, & Brown 1979, Easterbrook 1959, Combs and Taylor 1952, Baddeley 1981, Bacon 1974, Keinan 2002), developing an automated response will ensure that performing the task will not require more memory resources and the individual has resistance to the stress effects with respect to performance (Kivimaeki, & Lusa, 1994, Li, Baker, Grabowski, and Rebok 2001).

The second training method focuses on exposing the learner to stressful environments to gain familiarity with them, as well as enable the learner to build skills that would promote effective performance under stress. Research has shown that normal training procedures (training conducted under normal, non-stress conditions) do not avoid degradation of task performance when the task has to be performed under stress (Zakay & Wooler, 1984).This suggests that, the transfer of learning from training to operational conditions is affected if the training setup does

51

not approximate the actual environment. For our study, rather than recreate the surgical environment which is a complex set-up, and would still fall short of the real thing without a patient and the team which would exist in the operating room, we will study whether the coping mechanism developed during training under one type of stressor can be transferred when subjected to a novel stress environment. If this is true, we could greatly simplify the stressor for the MIS training environment, making it easy to implement, while still producing effective results in the operating room.

The three different training conditions designed for this study are:

Method 1: Training under non stressful conditions until performance goals are achieved

Method 2: Training under non stressful conditions until performance goals and the task has been "automated" (i.e., low workload has been achieved); and

Method 3: Training under a stressor until performance goals are achieved.

4.2 Background

4.2.1 Stressors in the Surgical Environment and their Impact on Surgeons

Surgical environments are highly stressful for surgeons due to a variety of reasons:

1. Surgical errors

2. Complexity of the task

3. Time pressure

4. Multitasking requirements

52

5. Equipment problems

6. Interpersonal issues

7. Distractions like noise

8. Personal factors like tiredness, hunger, illness, and physical discomfort. (Wetzel et al 2005).

Stress induced on surgeons has been found to be a major factor for performance deterioration in surgical environments by various studies (Wetzel et al., 2006, Van Gammert and Van Galen, 1997, Van Galen, Van Doorn, and Schomaker 1990).This is even more pronounced for surgeons who perform MIS surgeries as MIS is inherently more difficult to perform than open surgery (Berguer et al 2001). Participant performance in simulated minimally invasive surgery (MIS) tasks when under stress showed that stress can lead to impaired dexterity and an increase in psychomotor errors (Moorthy et al., 2003). In another study that aimed to understand the effects of realistic working conditions on surgical performance on novices and experts (Hsu et al. 2008) showed similar deterioration in laparoscopic performances. Pluyter and his colleagues (2010) demonstrated a decline in task score and an increase in task errors and operating time when a laparoscopic task was performed under distracting conditions.

The impact of stress on performance has become a primary concern in the military aviation (Prince et al 1994) and other medical areas in which effective performance under stress is mandatory. In these areas, various strategies have been attempted to mitigate the effect of stress.   Methods for coping for stress can be broadly classified into problem focused and emotion focused response training (Lazarus and Folkman, 1984, Lazarus and Launier, 1978).

Problem focused coping methods focus on minimizing or modifying the negative effect of stressors through cognitive processes. This involves providing information on the features of threatening situations and solutions to specific problems faced by the individual. These include training methods that focus on improving decision making and communication in the stressful environment, which leads the subject to develop confidence to overcome the negative effect of stress. This coping system has been found to be effective in controllable situations where the learners can manipulate stressors (Collins, Baum and Singer, 1983, Kaloupek and Stoupakis, 1985, Kaloupek, White and Wong 1984)

Emotion focused coping strategies work towards eliminating the distressful emotional component triggered by the stressor. In this type of training, the individual is taught to regulate the stressful effect rather than focus on specific problems that might have caused the stress. Emotion based coping include: controlling physiological responses through the use of biofeedback to individuals. Physiological control strategies alleviate the negative physiological reactions to stress by training people to respond in the same way as effective performers under stress conditions: calmness, relaxation, and control. Emotion focused coping has been found to be effective when the control over the stressful environment is limited or low (Strentz and Auerbach, 1988).

4.2.3 Different approaches to training for stress

In this section we will discuss the various training techniques that have been developed based on the aforementioned coping techniques.

4.2.3.1 Stress Training

The primary objective of stress training is to improve performance under stressful conditions by emotion coping strategies. The main emphasis of stress training is on relaxation or reduction of the stress response.   Various research studies have resulted in the development of different training techniques for handling stress, such as the use of biofeedback or stress inoculation training. Biofeedback is a means to monitor and controlling stress on an individual (Beatty and Legewie, 1977). For the stress training to be effective it is very important that the various indicators of stress used as biofeedback are reliable and valid. Although we considered examining this training, we found it difficult to obtain reliable physiological indicators to perform biofeedback (see Chapter 3).

Stress Inoculation training consists of an educational, a rehearsal and an application stage. Individuals are taught about the different stress management techniques such as cognitive restructuring, systematic desensitization, progressive relaxation in order to overcome stress (Wertkin, 1985). For example, in one study conducted by in a military setting, mental training techniques such as relaxation, mediation and imagery rehearsal were used as ways to manage stress.

4.2.3.2 Skill Training

The second type of technique for handling stress is based on skills training techniques. In skills training, the focus is on increasing the durability or automaticity of the skill. This is achieved by repeating a task over long periods of time. Well rehearsed tasks are found to be less influenced to stress thus preventing degradation in task performance. This allows the individual to enhance their sense of predictability and control of the task (Logan, 1985). Automaticity also reduces the

55

cognitive workload, which has been found to increase the speed and accuracy of performance (Wickens, 1984, Driskell et al 1992). In addition, decreasing the cognitive workload is important as the effect of stress reduces the available working memory resources even further (Mihake and Shah, 1999), which can be problematic if the cognitive workload is high. Hence, skill training that allows for automating the task has been found to ameliorate the effects of stress by producing an over learned behavior (Zajonic, 1965) and lower cognitive load.

4.2.3.3 Stress Exposure Training

Stress exposure training is another technique that has been found to improve performance under stress. Stress exposure training (SET) consists of two components: (1) coping with stress by using cognitive or behavioral strategies and (2) instructional design. The coping component includes the development of skills that reduce potential cognitive and psychomotor performance deficiencies resulting from specific stressors (Driskell and Salas, 1991, Hall, Driskell, Salas and Canon and-Bowers, 1992). The instructional design component involves gradual exposure to realistic stressors that enhance learning of coping skills and is based on basic principles of training design for skills acquisition (Meichenbaum, 1985, Smith, 1980). The primary basis of this training method lies in the findings from various research studies that training interventions that include task specific stressors have been successful in improving performance (Larsson, 1987, Meichenbaum, 1985, Novaco, 1988).

The objective of SET is to build skills that promote performance under stress, to build confidence and to enhance familiarity with the stress environment (Driskell, Hughes, Hall and Salas, 1992). SET has been found to improve performance and attitudes of individuals (Crocker et al 1988). For developing effective SET models it is important to develop training

environments which are realistic to that of the actual setting.   For our problem of MIS training, as it is difficult to replicate the stressors in the operating room in a simulated environment, we will examine whether any stressor can be used; namely; whether coping mechanism developed under one type of stressor can transfer to a novel stressor for a MIS task. This has been motivated by the study of Driskell and his colleagues (2001) which found that the beneficial effects of stress training can be retained when performed under novel stressors for non-MIS tasks. One of the experimental tasks was a spatial orientation task (Fitts, Weinstein, Rappaport, Anderson and Leonard, 1956) which required participants to compare a reference graph with a second set of graphs rotated to 90 or 270 degrees. Then they were to select the correct match. Another task was the Sternberg memory search task (Sternberg, 1969) which required participants to view a set of four letters and, once removed from view, the task required the participants to select whether a letter belonged to the memory set.

4.3 Methods

4.3.1 Task Objective

Understanding the conceptual relationship between the camera angle, the video view and the patient's anatomy is a fundamental problem in MIS. The primary objective of the task to be used is to train learners to understand the complex relationship that exists between the different camera angles and port locations with viewing a target in space.

4.3.2 Setup

 The general experimental set-up consists of a box trainer, which contains a simulated trocar (made of rubber) mounted on an x-y stage that can be placed and locked down in multiple positions (Section 2.6.1). To enable the learner to switch between different camera angles easily,

a simulated laparoscope was made based on a modified Endo Stick (3-D Med). The camera is mounted on a hinge inside a tube, where the angle of the hinge can be adjusted at the handle (between -45 to 45 degrees).

The basic task created was a camera navigation task, which was required to be performed at different ports and different camera angles. In a similar manner to (Eyal and Tendick, 2001) in virtual reality, we used long cylindrical boxes suspended in space in different positions and orientations for training the subjects– only in our case they are physical boxes mounted on a tree like structure. Each box contained a letter placed on a circular face inside the box at varying depths. The position and orientation of the boxes were varied so that, for some, the inside circular faces could be viewed only at a particular laparoscope position (left, center or right) and with a particular laparoscope lens. Other boxes were capable of being seen from more than one combination of position/lens angle.

In designing the "trees" that held the cylindrical boxes, the total number of targets that could be identified in each tree, and the number of targets that could be found in each camera and port combination, were varied to prevent any bias. The difficulty of each target was set using the geometric metric described in Section 3.3.1.2; the difficulty was chosen so that the overall difficulty of each of the task trees used was about equal. Each task tree was mounted on a task plate that could be slid into the training box easily. Five trees with on average of 18 targets were created for the training task. Four trees with on average 8 targets were created for the testing task. During both the training and testing phases we recorded performance measures such as number of major collisions using an accelerometer and position data using Motion Monitor.

The collisions were recorded using an accelerometer based collision detection system. A 3-axis accelerometer (Analog Devices ADXL 203) was placed inside a task tree trunk at half

58

the height of the trunk as shown in (Figure 2.14). A rubber sheet (Rubbercal) was placed between the task tree and the bottom of the trainer box to act as a damper to prevent "knocking on the box" being recorded as a collision by the accelerometer. A Labview program was used for collecting the real time data. The collection program assigned each collision as either major (>25gf) or minor (5-25gf) based on the intensity. It then displayed the results in real-time as the experiment was progressing.

In this study we recorded the various motion metrics using a Motion Monitor System. Motion monitor systems are data acquisition, analysis and visualization system developed by Innsport. The system used consisted of electromagnetic tracking devices, Ascensions mini-birds, consisting of 4 transmitters and a receiver. One transmitter was mounted on the tool tip and one was used to record the static 0 point. The sensors were sampled at 100 samples per second. The raw values were then processed by specialized software integrated with the system. The values were first exported and a Matlab algorithm which was used to calculate the path length details.

To aid in forming a conceptualization of the mechanics of viewing the task space, a target identification sheet provided to the learner so that they may keep track of what alphabet letters they viewed under which condition (i.e., position/lens angle).

4.3.2.1 Cognitive Load measurement: the NASA TLX

The cognitive load of a subject was measured using the NASA task load index NASA TLX (Hart and Staveland, 1988). It is a subjective workload assessment tool that allows users to perform assessment of their mental load while performing on human machine systems. The workload measures include: mental demand, physical demand, time pressure demand, performance, and effort and frustration level. All of these measures are applicable to the proposed camera navigation task. The users are required to rate their level of demand for each of the six

dimensions described previously on a twenty step bipolar scale with each step having five quality points for a maximum of 100 rating points. To arrive at a global score, a weighting procedure in which pair wise comparisons of each of the workload measures was carried out with all six dimensions by the participant. The number of times a dimension is chosen as more relevant than a pair is the weighting of that dimensional scale for a given task for the user. A workload score of 0 to 100 is obtained for each task by the individual rating for each dimension multiplied by its weighting factor, then summing across scales and dividing by 15 (the number of pair wise comparisons). A value closer to 0 indicates that the user requires low cognitive resources to perform the task.

The sensitivity of the NASA TLX has been demonstrated in a wide range of tasks such as real (Shively et al 1987) and simulated flight tasks (Battiste and Bortolussi, 1988; Corwin, Sandry-Garza, Biferno, Boucek, Logan, Jonsson, & Metalis, 1989; Nataupsky & Abbott, 1987; Tsang & Johnson, 1989; Vidulich & Bortolussi, 1988), using remote control vehicles (Byers et al 1989) and in other laboratory research programs (Hart and Staveland, 1988). Various studies have found NASA TLX to be sensitive, having high convergent and concurrent validity, and diagnosticity, as compared to other techniques, such as the subjective assessment technique (SWAT, Reid and Nygren, 1988) and workload profile WP (Wickens 1984) methods. The NASA TLX was administered during Training for subjects in Group 2 alone each time they reach the performance measure of finding greater than 80% targets in their attempt while having zero major collision. While in the testing phase all the subjects were required to take the test after completion of each testing tree.

60

Figure 4.1: Laparoscopic Trainer system

4.3.3 Experimental Design

4.3.3.1 Participants

A total of 48 subjects (males (47.9%) and females (52.1%)) participated in the study. Eleven subjects were students from the School of Medicine and the remaining was from the School of Engineering and other departments. Before the study began, participants were asked whether they had previous gaming experience (and how much), whether their dominant hand was right or left, and whether they had previous laparoscopic experience. The sex of the participant was also recorded. All the subjects either worked or studied at Virginia Commonwealth University.

4.3.3.2 Training

All subjects were randomly assigned to one of three groups:

 Group 1: Subjects trained under no stress until high performance scores were attained,

Group 2: Subjects trained under no stress until high performance and low cognitive load were attained and

Group 3: Subjects trained under stress until high performance scores were attained.

The basic learning task considered for this study required the subjects to learn the use of laparoscopes with 0 and 30 degree angles at two of the most standard locations within the operating field (to the left of and in the center of where the laparoscopic tools are placed). The specific task was for learners to find the targets inside the cylinders presented in a sequence of task trees until they reached the stopping condition for their assigned group. Five different task trees were available for training. They were presented in random order and, if all five were used, the trees would be presented in a second block in random order.

4.3.3.2.1Training Instructions

On the first day, learners were instructed on the way the laparoscope functions and the basics of laparoscopic surgery. On both training days (if the second one was needed),they were provided with 15 minutes of practice time to familiarize themselves with the system before any task trees were placed in the trainer box. They were also told that the basic task for them to perform during learning is to locate letters inside cylinders placed inside the training box at different positions and orientations. Then they were told that each tree can be accessed using four different combinations of camera angle and port, and that they should follow the same order of camera angle and port location, given by the experimenter, for all the trees. This was: 0-center, 30–center, 0-left and 30-left. They were also instructed to avoid making any major collisions, where a major collision was defined when the force of contact is greater than 25gF.

62

Participants were also asked to create a mental picture of the relationship between the laparoscope in the different positions and lens angles, and the overlapping field of views that this produces. They were told that they will be provided with a target identification sheet to aid with this task. It was explained to participants that, in the operating room, it was required to make a decision about both the port position and lens angle before the operation is performed; any wrong selection would result in a severe penalty requiring another opening in the patient, and a noticeable one for changing lens angle, as this would require switching laparoscopes. However, they were told that during learning there is no such penalty, except that it is of benefit that they should construct an appropriate mental model to be able to know what to do in the operating room. This will be validated in the testing task. There was no time limit for any trial.

4.3.3.2.2 Duration of Task

Each learner was trained for a maximum of two days to learn the task or until the stopping criteria was met, whichever was shorter. The two days were required to be within a week of each other. For each day, there was an upper limit to the training of 3 hours.

4.3.3.3 Stopping Conditions

Group 1

The training condition for this group consisted of the basic camera navigation and target identification task until the required performance level was reached. They were not subjected to any external stressor. The stopping condition for this group was to at-least find 80% targets with zero major collisions.

63

Group 2

The training objective of this group was the requirement for the learners to attain automaticity of the task being learnt, where automaticity was defined as reaching low-cognitive load. The training condition for this group consisted of the basic camera navigation and target identification task until the required high performance and low cognitive levels were reached. They were not subjected to any external stressor. The stopping condition for this group was set to at-least find 80% targets, have zero major collisions and low cognitive load calculated (as determined by the NASA Task Load Index score.) "Low" cognitive load was determined as an average of 20 across measures. The cognitive load was measured at the end of each trial whenever the percentage targets found was greater than 80%

Group 3

The training condition for this group consisted of the basic camera navigation and target identification task until the required performance level was reached. They were subjected to an external stressor –arm stretch which required the subjects to hold their arms at shoulder height for the entire duration of the trial. The stopping condition for this group was to at-least find 80% targets with zero major collisions.

4.3.3.4 Testing Task.

The testing phase trials were conducted after seven days from the last day of training. This was chosen as it has been found that rest periods after training ensures better retention and transfer of learnt skill (Pierre Maquet et al, 2001). The interval period after training is found to aid consolidation of the learned task. Evidence indicates improvement in performance on a task when tested after several days of training (Karni and Sagi 1993; Karni et al. 1995).

64

As our objective is to have learners develop a mental model to enable the transfer of their knowledge to new problems and the O.R., an ideal testing task would be to examine performance in the O.R. over several different operations after training. However as it is impractical to achieve this, the transfer test developed for this study will assess the ability of the learner to decide on the camera angle and port location configurations for a task setup. This test simulates the decision making process and manipulations a surgeon is subjected to in the O.R. The specific task to be considered is:

 (1) Choosing the position of the laparoscope port and angle of the lens to be used based on 3D images provided, and

(2) Manipulating the laparoscope to identify the targets located in the workspace.
The subjects were presented with a 3D picture containing three views (front, left and right side) of the task tree and were instructed to select the appropriate camera angle (0 or 30 degree) and port (left or center) for viewing the targets.

As our objective was to have learners develop a mental model to enable the transfer of their knowledge to new problems and the O.R., the transfer task was used to evaluate the learners' performance with a novel stressor (i.e., one not used before in any training condition). All the subjects were subjected to a neuro-muscular electrical stimulator (NMES) while performing the target identification task. Different stressors were used during the training and the testing phase to ensure that learners in Group 3 were not adapted to the particular stressor which would result in unfair advantage for them.

The testing phase consisted of four task trees with eight targets in two and seven targets in the other two trees. The primary evaluation was the number of letters participants found.

65

Secondary variables used for evaluation were: (a) the number of errors they performed (i.e., making undesirable contact with the laparoscope),(b) the amount of time taken and motion metrics such as (c) Path length, (d)Smoothness of motion,(e)Selection of port and camera angle,(f)Derived measures :Path length per target, Path length per time, Path length per collision and Collision per target. They had a time limit of one and half hours for performing all fours tests. The NASA TLX was administered after each of the four tests.
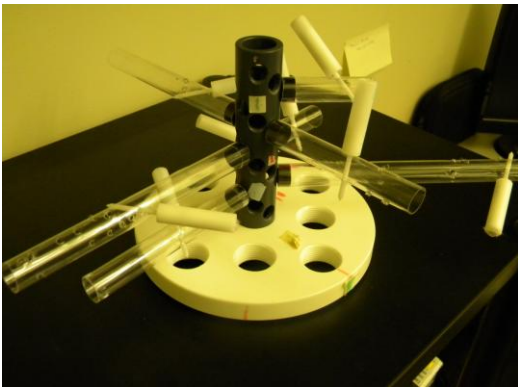
Figure 4.2 Left side view
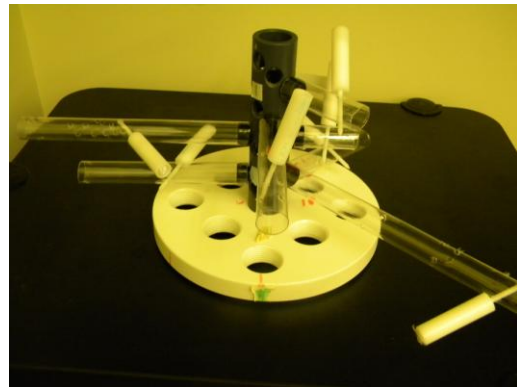
Figure 4.3 Front side view
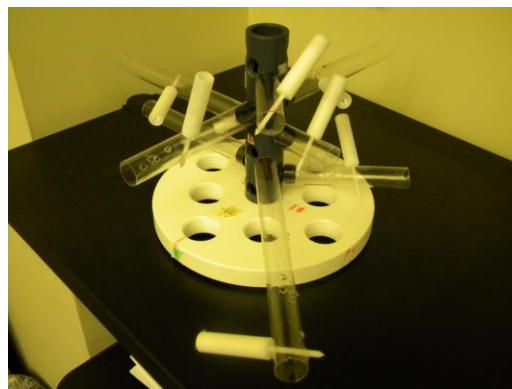
Figure 4.4 Right side view

4.3.3.5 Calculation of Variables

The different metrics in this study typically assess the learner's proficiency level by means of recording

- Total number of targets identified

- Path length traversed by the instrument over time: Recorded using Motion monitor.

- Smoothness of motion: Calculated by the third time derivate of position data recorded using Motion Monitor

- Number of collisions: Recorded using the accelerometer based collision detection system ADXL 203

- Time to completion(in min)

- Derived measures:

  a. Time per targets : Total time for completion for each trial divided by total number of targets identified for each trial

  b. Collision per target: Total number of collisions for each trial divided by total number of targets identified for each trial

  c. Path length per Target: Total path length taken for each trial divided by total number of targets identified for each trial

  d. Path length per collisions: Total path length taken for each trial divided by total number of collisions for each trial

67

- Selection of Camera: Number of correct camera angle selection

- Selection of Port: Number of correct port selection

Generally a comparative analysis is performed on the results of one or more of these measures at the end of training to determine the extent of learning.

4.3.3.5.1 Calculation of motion metrics

Currently the motion metrics are not fed back to the user while performing the task but are used for assessment purposes in terms of examining testing performance. In this study, the position data was then transformed mathematically to determine the path length of the tool tip and smoothness of motion values.

The total path length in 3D space was found by using the distance formula between two points in 3D space (x1, y1, z1) and (x2, y2, z2) and summing the individual distance of consecutive points over the entire duration of the trial.

D1 = √ ((x2-x1) ^2 + (y2-y1) ^2 + (z2-z1) ^2)

Total distance or Path length =∑D1, D2….D n where n is the last point recorded.

Smoothness of motion: This is calculated by performing a third time-derivative of position in cm/s^3.

$$j = \sqrt{(\frac{d^3x}{dt^3})^2 + (\frac{d^3y}{dt^3})^2 + (\frac{d^3z}{dt^3})^2}$$

Motion smoothness is derived from the integrated squared jerk as

68

$$J = \sqrt{\frac{1}{2} \int_0^T j^2 dt}.$$

Since for this study established motion metrics were used for assessment no specific tests were required to test their accuracy or reliablity.

4.3.3.6 Statistical Methods

The study included 48 participants: 23 participants were male, 47 were right hand dominant, 7 had laparoscopy experience and 24 had gaming experience. We classified our data based on their type. The nominal data for this study were the sex of the subject, previous laparoscopic experience, and previous gaming experience and counts and percentages were used for the analysis of them. Similarly means, medians, standard deviations and interquartile ranges were used for analysis of the continuous data namely training time and trials attempted by the subjects. Specific Generalized linear mixed models (GLMM) were developed for the different performance variables to determine the effectiveness of the three training methods. Since GLMM model is unique and complex for each outcome measure they are thus described in more detail in the following sections

4.4 Subjective characteristics

The subjective characteristics for the sample of 48 participants are summarized in Table 4.1. The sample had similar rates of males (47.9%) and females (52.1%) and participants with and without gaming experience (50%), were primarily right handed (98%), and most of them had no previous laparoscopy experience (85%).

Table 4.1: Subject Characteristics

Subject Characteristics		Count	Percent
Sex	Male	23	47.9
	Female	25	52.1
Dominant Hand	Right	47	97.9
	Left	1	2.1
Laparoscopy Experience	Yes	7	14.6
	No	41	85.4
Gaming Experience	Yes	24	50
	No	24	50

4.5 Training characteristics between groups

We observed that both the training characteristics i.e. the total time taken $F (2, 45) = 26.6$, $p < 0.0001$) and total number of training trials attempted (chi-square $(2) = 16.15$, $p = 0.0003$) were significantly different between the three groups. Participants in Group 2 took significantly longer training time and more training trials than Group 1 or Group 3 .The primary reason for this difference is that the subjects in this group were required to continue training until they reach the required performance criterion which additionally required a low cognitive load of less than 20 on the NASA TLX score index while those subjects in Group 1 and Group 3 were only limited by the performance goal. The total training time and the total number of trial attempts is summarized by group in Table 4.2

Table 4.2: Training Characteristics

	Total Training Time (minutes)		Total Number of Trials Attempted	
	Mean	SD	Median	IQR
Group 1	129.4	60.7	5.5	3-7
Group 2	274.6	91.8	11	8-14.75
Group 3	115.2	44.3	4	3-6.25

4.6 Training differences by Gaming experience

Comparing subjects based on their previous gaming experience (Table 4.3) indicated that there was no significant difference in either the training time (t (46) = 0.30, p = 0.7674) or the number of training trials (chi-square (1) = 0.03, p = 0.8597) attempted.

Table 4.3Gaming Characteristics

	Total Training Time (minutes)		Total Number of Trials Attempted	
	Mean	SD	Median	IQR
Previous Gaming Experience	177.4	90.4	6	3.25-9.75
No Previous Gaming Experience	168.8	108.6	7	4-8.75

4.7 Primary performance measures

4.7.1 Proportion of Targets Identified

4.7.1.1 Statistical method

The proportion of targets identified was modeled at the target level with a generalized linear mixed effects model (GLMM) assuming a binary distribution for the response and a logit link function ($X\beta=\ln(\mu/1-\mu)$) between the response (found vs. not found) and the model effects. This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which groups were significantly better for the specific performance measure analyzed. Effects were interpreted using odds ratios describing how many times greater the odds of identifying a target (versus not identifying a target) was for one group versus another group, where the odds are defined as the probability of identifying a target divided by the probability of not identifying a target.

4.7.1.2 Results

The GLMM indicated that there were significant differences in the proportion of targets identified between the groups (p=0.0140), after controlling for tree (p < 0.0001), previous laparoscopy experience (p = 0.0070), camera selection (p = 0.0010), and level of target difficulty (p < 0.0001). The odds of identifying a target were 65.9% (or 1.659 times) greater for group 2 versus group 1 (p = 0.0066) and 53.4% (or 1.534 times) greater for group 2 versus group 3 (p = 0.0199); the odds were not significantly different between groups 1 and 3 (p = 0.6599) .There was no evidence of significant interaction effects for group by tree (p = 0.5928), number of training trials by group (p = 0.1951), or total training time by group (p = 0.3620), or effects due to number of training trials (p = 0.9116), total training time (p = 0.7041), sex (p = 0.2311), previous gaming experience (p = 0.1908), or port selection (p = 0.1286).

The proportion of targets identified for each of the groups is plotted below (Figure 4.5) and is summarized in Table 4.4 along with the odds of identifying the targets versus not identifying the targets.

73
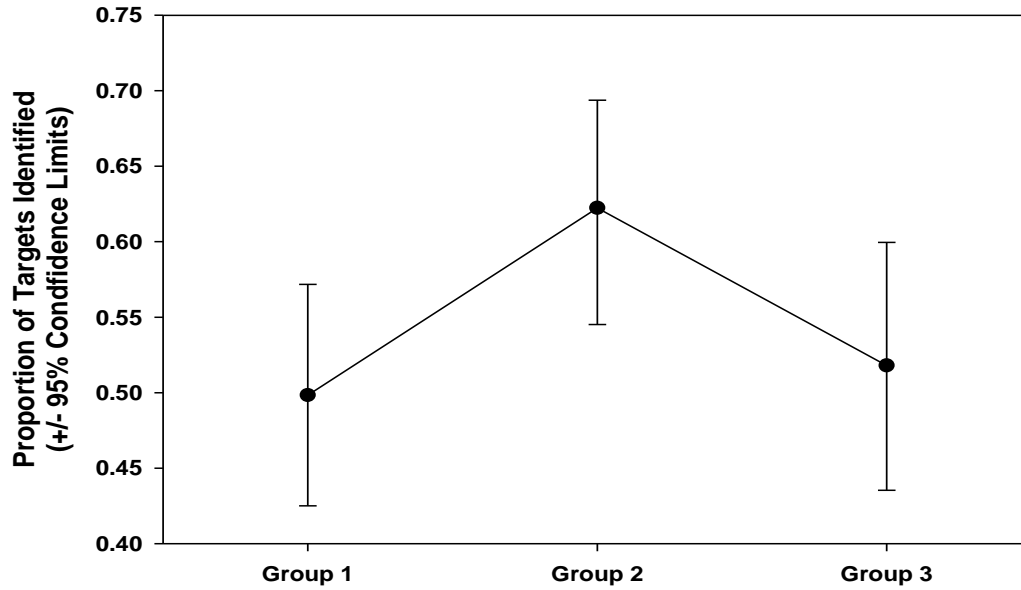
Figure 4.5: Proportion of Targets Identified by Group

Table 4.4: Proportion of Targets Identified and Odds of Identifying Targets versus not
Identifying Targets by Group

Group	Proportion of Targets Identified			Odds Ratio	
	Estimate	SE	95% CI	Estimate	95% CI
1	0.4984	0.0368	(0.4251, 0.5718)	0.994	(0.739, 1.336)
2	0.6224	0.0374	(0.5452, 0.6938)	1.648	(1.199, 2.266)
3	0.5180	0.0413	(0.4354, 0.5996)	1.075	(0.771, 1.497)

SE = standard error; CI = confidence interval

The proportion of targets identified and the odds ratio based on the laparoscopy experience of the
subject is summarized below in Table 4.5 and Table 4.6. The odds of identifying the target were
significantly greater for those with previous laparoscopy experience versus those without (p =
0.0070).

Table 4.5: Proportion of Targets Identified and Odds of Identifying Targets versus not

Identifying Targets by Previous Laparoscopy Experience

Laparoscopy Experience	Proportion of targets identified			Odds Ratio	
	Estimate	SE	95% CI	Estimate	95% CI
No	0.4717	0.0228	(0.4266, 0.5173)	0.893	(0.744, 1.072)
Yes	0.6201	0.0490	(0.5183, 0.7124)	1.633	(1.076, 2.477)

SE = standard error; CI = confidence interval

Table 4.6: Odds Ratios for Identifying Targets by Previous Laparoscopy Experience

Tree	OR	95% CI	p-value
Laparoscopy Experience			
Yes vs. No	1.828	(1.189, 2.812)	0.0070

OR = odds ratio; CI = confidence interval

The odds of identifying a target were significantly greater for difficulty level 1 versus levels 2 (p < 0.0001) and 3 (p < 0.0001), and for difficulty level 3 versus level 2 (p = 0.0489). The proportion of targets identified and the odds ratio for target difficulty is summarized in Table 4.7, 4.8 below.

Table 4.7: Proportion of Targets Identified and Odds of Identifying Targets versus not

Identifying Targets by Target Difficulty

Target Difficulty	Proportion of targets identified			Odds Ratio	
	Estimate	SE	95% CI	Estimate	95% CI
1	0.6666	0.0325	(0.5995, 0.7276)	2.000	(1.497, 2.671)
2	0.4506	0.0353	(0.3823, 0.5208)	0.820	(0.619, 1.087)
3	0.5176	0.0356	(0.4473, 0.5873)	1.073	(0.809, 1.423)

SE = standard error; CI = confidence interval

75

Table 4.8: Odds Ratios for Identifying Targets comparing Target Difficulty Levels

Tree	OR	95% CI	p-value
Target Difficulty			
1 vs. 2	2.438	(1.856, 3.203)	< 0.0001
1 vs. 3	1.864	(1.422, 2.442)	< 0.0001
3 vs. 2	1.308	(1.001, 1.709)	0.0489

OR = odds ratio; CI = confidence interval

The odds of identifying a target were significantly greater for camera angle 30º versus 0º (p <

0.0001) is summarized in Table 4.9 and 4.10 below.

Table 4.9: Proportion of Targets Identified and Odds of Identifying Targets versus not

Identifying Targets by Camera angle selected

	Proportion of targets identified			Odds Ratio	
Camera Selected	Estimate	SE	95% CI	Estimate	95% CI
0º	0.4819	0.0428	(0.3987, 0.5661)	0.930	(0.663, 1.304)
30 º	0.6105	0.0259	(0.5575, 0.6610)	1.567	(1.260, 1.950)

SE = standard error; CI = confidence interval

Table 4.10: Odds Ratios for Identifying Targets Comparing the Trees, Previous Laparoscopy

Experience Groups, and Target Difficulty Levels

Tree	OR	95% CI	p-value
Camera Selection			
30º vs. 0º	1.685	(1.236, 2.298)	0.0010

OR = odds ratio; CI = confidence interval

The odds of identifying a target were significantly greater for tree 2 versus trees 3 (p = 0.0164)

and 4 (p < 0.0001), for tree 1 versus 4 (p < 0.0001), and tree 3 versus 4 (p < 0.0001); there was

no significant difference between trees 1 and 2 (p = 0.0694) or between trees 1 and 3 (p =

0.4796).

76

Table 4.11: Proportion of Targets Identified and Odds of Identifying Targets versus not

Identifying Targets by Tree selected

	Proportion			Odds	
Tree	Estimate	SE	95% CI	Estimate	95% CI
1	0.5929	0.0373	(0.5178, 0.6639)	1.456	(1.074, 1.975)
2	0.6600	0.0371	(0.5836, 0.7289)	1.941	(1.402, 2.689)
3	0.5656	0.0367	(0.4922, 0.6362)	1.302	(0.969, 1.749)
4	0.3660	0.0347	(0.3007, 0.4368)	0.577	(0.430, 0.776)

Table 4.12: Odds Ratios for Identifying Targets Comparing the Trees

Tree	OR	95% CI	p-value
2 vs. 1	1.333	(0.977, 1.818)	0.0694
2 vs. 3	1.491	(1.076, 2.066)	0.0164
2 vs. 4	3.362	(2.441, 4.632)	< 0.0001
1 vs. 3	1.119	(0.820, 1.527)	0.4796
1 vs. 4	2.522	(1.863, 3.414)	< 0.0001
3 vs. 4	2.255	(1.653, 3.076)	< 0.0001

OR = odds ratio; CI = confidence interval

4.7.2 Number of Major Collisions per Tree

4.7.2.1 Statistical Method

The number of major collisions was modeled at the tree level with a generalized linear mixed

effects model (GLMM) assuming a Poisson distribution for the response (number of major

collisions) and a log link function ($X\beta = \ln (\mu)$) that provides the relationship between the linear

predictor (model effects)and the mean of the distribution function (responses). This model also

took into account the variations in the responses within as well as between subjects for each of

the performance measures. The main effects for group and tree and the interaction effects for

group by tree, training time by group and number of training attempts by group were included in

77

the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values $\geq$ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which groups were significantly better for the specific performance measure analyzed. Finally effects were interpreted using ratios of the means for each of the measure to determine how many times greater the mean of that measure for one group was compared to the other group. Effects were interpreted using ratios of the means describing how many times greater the mean number of major collisions was for one group versus another group.

4.7.2.2 Results

As one of the subjects in Group 1 made an unusually higher number of collisions he was considered an outlier and removed from the analysis. The GLMM indicated that the effect of group on the number of major collisions depends on the total number of training trials (p = 0.0300), after controlling for tree (p = 0.0037). There was no evidence of significant interaction effects for total training time by group (p = 0.7845) or group by tree (p = 0.7838) or effects due to camera selection (p = 0.8046), total training time (p = 0.9769), previous gaming experience (p = 0.5476), port selection (p = 0.4166), sex (p = 0.1848), or previous laparoscopy experience (p = 0.1836).

There was a significant interaction effect between group and number of training trials .Subjects in Group 1, who trained with 1 to 9 total number of trial attempts exhibited significant increases

in the number of major collisions with increases in training trial attempts (p = 0.0250); whereas subjects in group 3, training with 2 to 10 trial attempts, and group 2, training with 2 to 21 trial attempts, showed trends for the mean number of collisions to nominally decrease with increases in the number of training trial attempts. The decreasing trends for groups 2 (p = 0.5315) and 3 (p = 0.1818) were not significant however.
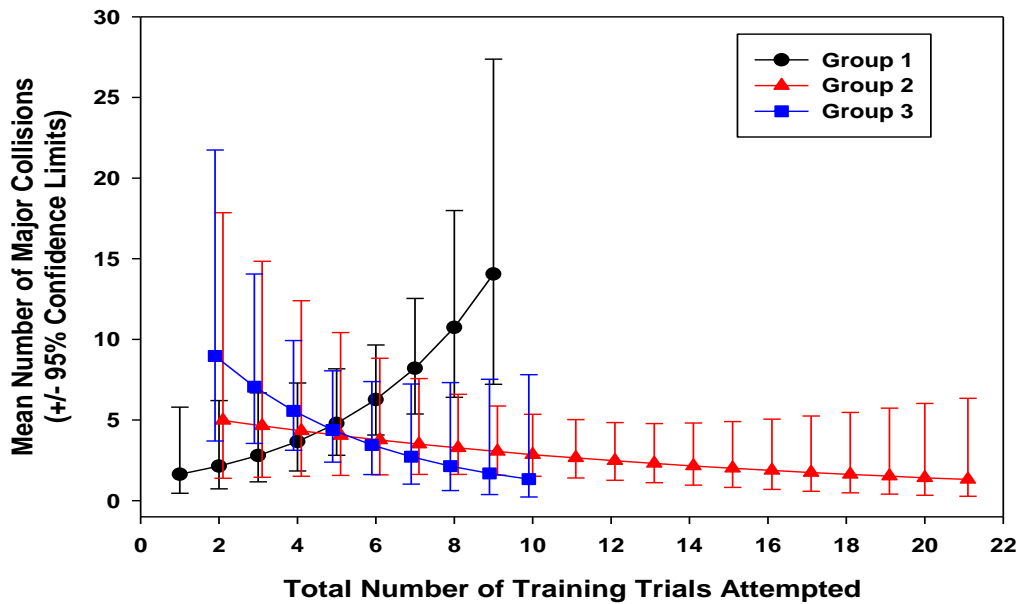
Figure 4.6: Number of Major Collisions by Group across Total Number of Training Trials Attempted by Group

The number of major collisions was significantly greater for those with previous laparoscopy experience versus those without (p = 0.0097).This observation is different from results from other literatures. The quality and the level of experience could possibly provide more insights on this variable on performance but a separate study is required to determine this.

Table 4.13: Ratios of Mean Number of Collisions comparing Previous Laparoscopy Experience

Groups

Laparoscopy Experience	Ratio	95% CI	p-value
Yes vs. No	2.510	(1.265, 4.978)	0.0097

CI = confidence interval

Table 4.14: Mean Number of Major Collisions by Previous Laparoscopy Experience

	Mean		
Laparoscopy Experience			
No	2.7008	0.6346	(1.6826, 4.3352)
Yes	6.7778	2.2013	(3.5218, 13.0438)

SE = standard error; CI = confidence interval

Table 4.15: Mean Number of Major Collisions by Tree (Excluding Subject ID #1)

	Mean		
Tree	Estimate	SE	95% CI
1	3.6162	1.0075	(2.0703, 6.3165)
2	2.4517	0.5091	(1.6143, 3.7235)
3	1.7366	0.4330	(1.0542, 2.8606)
4	4.0044	1.0118	(2.4147, 6.6408)

SE = standard error; CI = confidence interval

Table 4.16: Ratios of Mean Number of Collisions Comparing the Trees (Excluding Subject

ID#1)

Tree	Ratio	95% CI	p-value
1 vs. 2	1.475	(0.939, 2.316)	0.0896
1 vs. 3	2.0824	(1.293, 3.354)	0.0033
1 vs. 4	0.903	(0.533, 1.529)	0.6985
4 vs. 2	1.633	(1.028, 2.596)	0.0384
4 vs. 3	2.306	(1.441, 3.691)	0.0008
2 vs. 3	1.412	(0.980, 2.035)	0.0639

CI = confidence interval

4.7.3 Total Time taken to Identify Targets

4.7.3.1 Statistical Method

The total time taken to identify targets was modeled at the tree level with a generalized linear

mixed effects model (GLMM) assuming a normal distribution for the response (number of major

collisions) and an identity link function ($X\beta = \mu$) that provides the relationship between the

linear predictor (model effects)and the mean of the distribution function (responses).This model

also took into account the variations in the responses within as well as between subjects for each

of the performance measures. The main effects for group and tree and the interaction effects for

group by tree, training time by group and number of training attempts by group were included in

the model. The fixed effects for previous gaming experience, previous laparoscopy experience,

sex, total training time, total number of training trials attempted, camera selected, port selected

were initially included in the model and later a backward selection process was used to remove

any of the effects except that of group and tree if they did not contribute in improving the model

fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed

between groups , specific contrasts were tested to determine which groups were significantly

81

better for the specific performance measure analyzed. Finally effects were interpreted using ratios of the means of time taken to identify the target to determine how many times greater the mean for one group is compared to the other group. Effects were interpreted using differences in the means describing how much longer (in minutes) the time to identify targets was for one group is versus another group.

4.7.3.2 Results

The GLMM indicated that there was no significant differences in the total time taken to identify targets between the groups ($p = 0.1694$), after controlling for tree ($p = 0.9087$), total training time ($p = 0.0016$), and total number of training trial attempts ($p = 0.0082$). There was also no evidence of significant interaction effects for group by tree ($p = 0.6916$), number of training trials by group ($p = 0.4731$), or total training time by group ($p = 0.8811$), or effects due to sex ($p = 0.4647$), previous gaming experience ($p = 0.3851$), previous laparoscopy experience ($p = 0.1112$), camera selection ($p = 0.7413$), or port selection ($p = 0.9299$).

The mean total time taken to identify targets for each of the groups is plotted in Figure 4.7 and is summarized in Table 4.17. The differences comparing the groups are summarized in Table 4.17. While group 2 nominally had a lower mean time (6.4 minutes) than groups 1 (9.0 minutes) and 3 (9.1 minutes), none of the pair wise comparisons were statistically significant (as evident by the test of the group main effect: $p = 0.1694$).
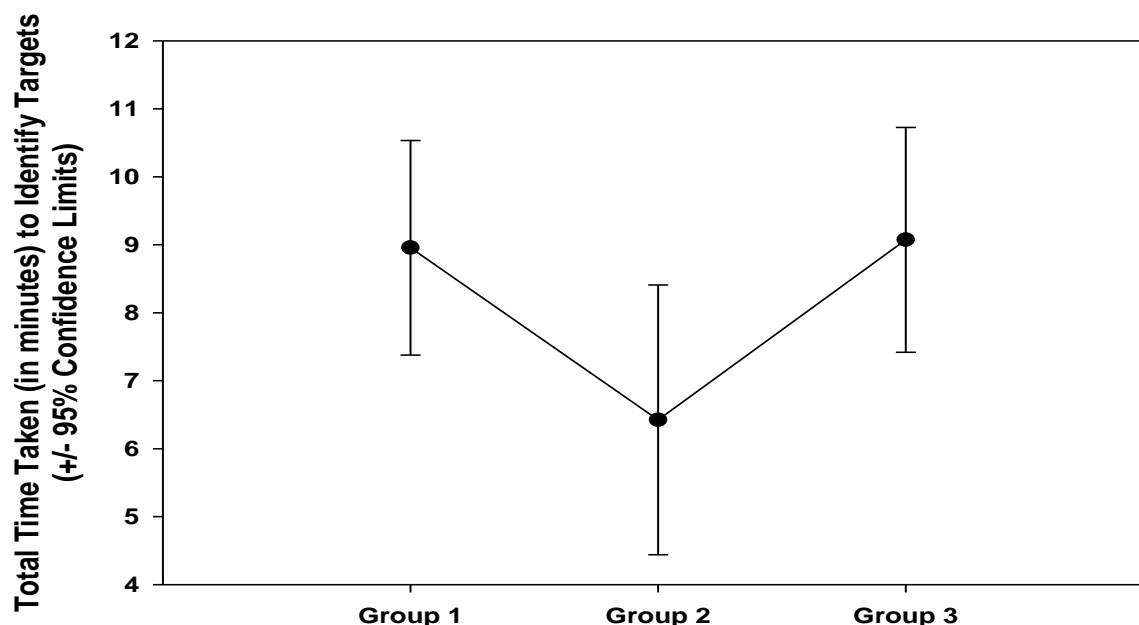
Figure 4.7: Total Time Taken to Identify Targets by Group

Table 4.17: Total Time (in minutes) Taken to Identify Targets by Group

Group	Mean	SE	95% CI
1	8.9564	0.7827	(7.3781, 10.5347)
2	6.4241	0.9842	(4.4393, 8.4088)
3	9.0727	0.8201	(7.4190, 10.7264)

SE = standard error; CI = confidence interval

The total time taken to identify targets was positively associated with the total training time (p = 0.0016) and negatively associated with the total number of trial attempts (p = 0.0082). A ten minute increase in the total training time was associated with a 0.2908 minute increase in the total time take to identify targets (SE = 0.0865, 95% CI = 0.1162, 0.4653).

For each additional trial attempted, the total time taken to identify targets decreased by 0.466 minutes (SE = 0.168, 95% CI = 0.1271, 0.8057).

83

4.7.4 Path Length

4.7.4.1 Statistical Methods

The total path length was modeled at the tree level with a generalized linear mixed effects model (GLMM) assuming a log-normal distribution ($\Xi\beta = \mu$) that provides the relationship between the linear predictor (model effects) and the mean of the distribution function (number of collisions). This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which groups were significantly better for the specific performance measure analyzed. If there was evidence of differences between the groups, specific contrasts were tested to determine which groups had significantly different mean path length. Effects were interpreted using ratios of the means describing how many times greater the mean path length was for one group versus another group.

4.7.4.2 Results

The GLMM indicated that there were not significant differences in the total path length between the groups (p = 0.1235), after controlling for tree (p = 0.1593), total training time (p = 0.0242),

84

and previous laparoscopy experience (p = 0.0117). There was also no evidence of significant interaction effects for group by tree (p = 0.8310), number of training trials by group (p = 0.8160), or total training time by group (p = 0.1749), or effects due to previous gaming experience (p = 0.7912), sex (p = 0.0732), total number of training trial attempts (p = 0.1355), camera selection (p = 0.0670), or port selection (p = 0.0532).

Group 2 nominally had a lower mean path length than groups 1 and 3, none of the pair wise comparisons were statistically significant (as evident by the test of the group main effect: p = 0.1235). The mean total path length for each of the groups is plotted in Figure 4.8 and is summarized in Table 4.18.

Table 4.18: Total Path Length (in mm) by Group

Group	Mean	95% CI
1	30.5489	(24.4283, 38.2031)
2	21.5522	(16.0286, 28.9793)
3	31.6461	(24.6927, 40.5576)

CI = confidence interval

85
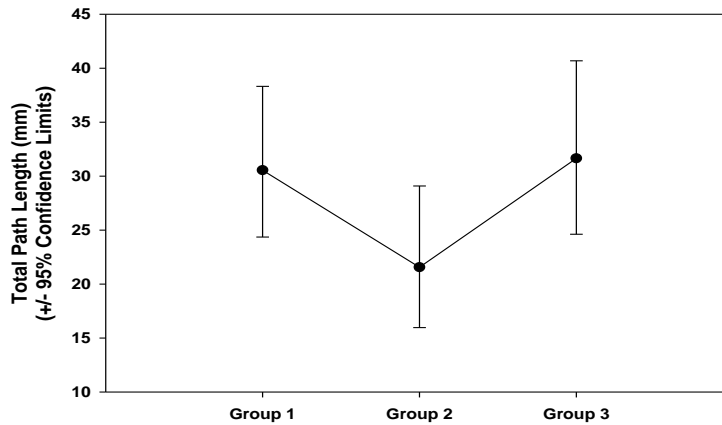
Figure 4.8: Mean Total Path Length (mm) by Group

There was a significant positive relationship between the total path length and the total time training (p = 0.0242).A 10 minute increase in the training time was associated with an increase in the log total path length of 0.2018 (95% CI = 0.0277, 0.3760). However Group 2 had lower path lengths than Group 1 and Group 3.
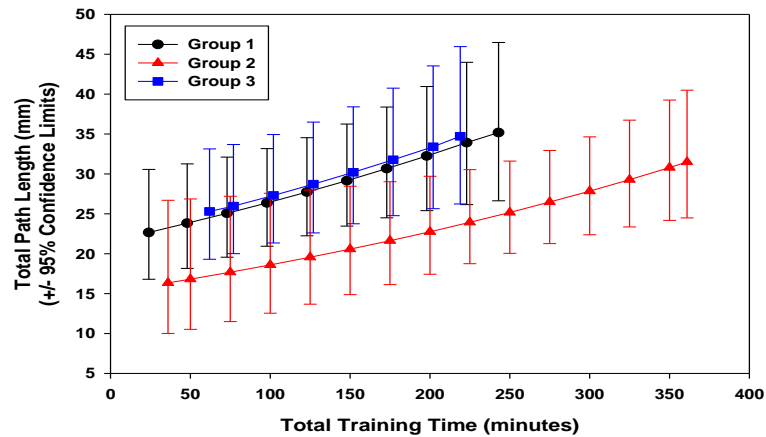
Figure 4.9: Ratios of the Mean Total Path Length per training trials

There was a significant difference in the mean number of major collisions between those with and without previous laparoscopy experience (p = 0.0117). The mean total path length for each

86

laparoscopy experience subgroup is summarized in Table 4.19.Those with previous experience had mean total path length s that were 1.5422 times larger than those without previous experience (95% CI = 1.1066, 2.1492).

Table 4.19: Mean Total Path Length (mm) by Previous Laparoscopy Experience

Laparoscopy Experience	Mean	95% CI
No	22.1583	(19.6610, 24.9728)
Yes	34.1719	(25.2214, 46.2987)

CI = confidence interval

4.7.5 Selection of Camera

4.7.5.1 Statistical Method

The probability of correct camera selection was modeled at the tree level with a generalized linear mixed effects model (GLMM) assuming a binary distribution for the response (correct camera selected) and a logit link function ($\Xi\beta=\lambda\nu(\mu/1-\mu)$) that provides the relationship between the linear predictor (model effects)and the mean of the distribution function (responses, Camera selection). This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which

groups were significantly better for the specific performance measure analyzed. If there was evidence of differences between the groups, specific contrasts were tested to determine which groups had significantly greater odds of correct camera selection. Effects were interpreted using odds ratios describing how many times greater the odds of correct camera selection (versus incorrect camera selection) was for one group versus another group, where the odds are defined as the probability of correct camera selection divided by the probability of incorrect camera selection.

4.7.5.2 Results

The GLMM indicated that the effect of group on the probability of selecting the correct camera angle depends on the total number of training trials ($p = 0.0068$) and on the total training time ($p = 0.0084$), after controlling for tree ($p < 0.0001$). There was no evidence of a significant interaction effect for group by tree ($p = 0.8283$) or effects due to previous gaming experience ($p = 0.6763$), previous laparoscopy experience ($p = 0.2726$) or sex ($p = 0.1730$).

The probability of correctly selecting the camera for each of the groups across the observed values of total number of training trials attempted is plotted in Figure 4.10. In general, subjects in Group 1 exhibited significant increases in the probability of correctly selecting the camera angle with increases in training trial attempts ($p = 0.0016$), subjects in group 2 showed trends for the probability of correctly selecting the camera to nominally decrease and group 3 showed trends for the probability of correctly selecting the camera to nominally increase, respectively, with increases in the number of training trial attempts. The trends for groups 2 ($p = 0.5310$) and 3 ($p = 0.5676$) were not significant however.

88
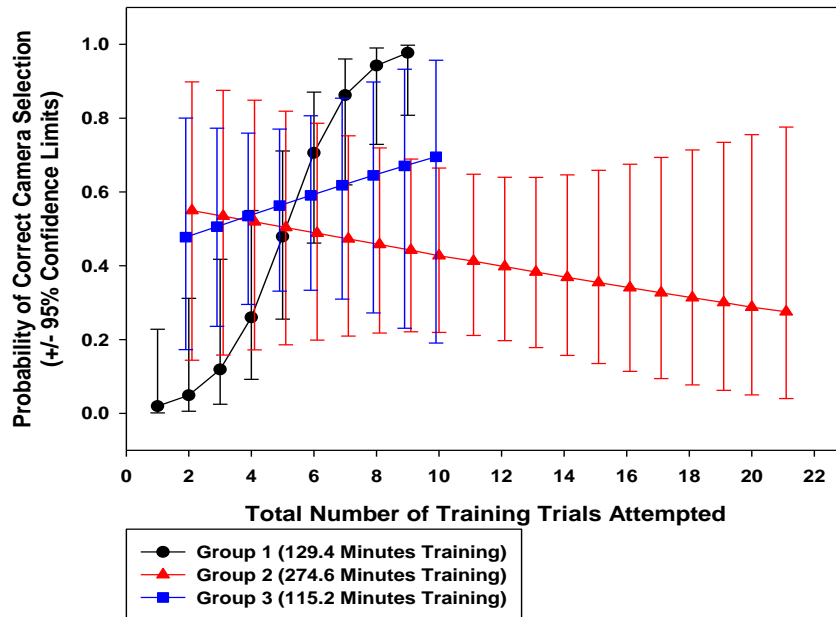
Figure 4.10: Probability of Correct Camera Selection by Group across Total Number of Training

Trials Attempted

The probability of correctly selecting the camera for each of the groups across the observed

values of total time training is plotted in Figure 4.11. In general, subjects in Group 1 exhibited

significant decreases in probability of correctly selecting the camera angle with increases in the

total training time (p = 0.0036) whereas subjects in group 2 and group 3 showed trends for the

probability of correctly selecting the camera to nominally increase and decrease, respectively,

with increases in the total training time. The trends for groups 2 (p = 0.4019) and 3 (p = 0.1155)

were not significant however.
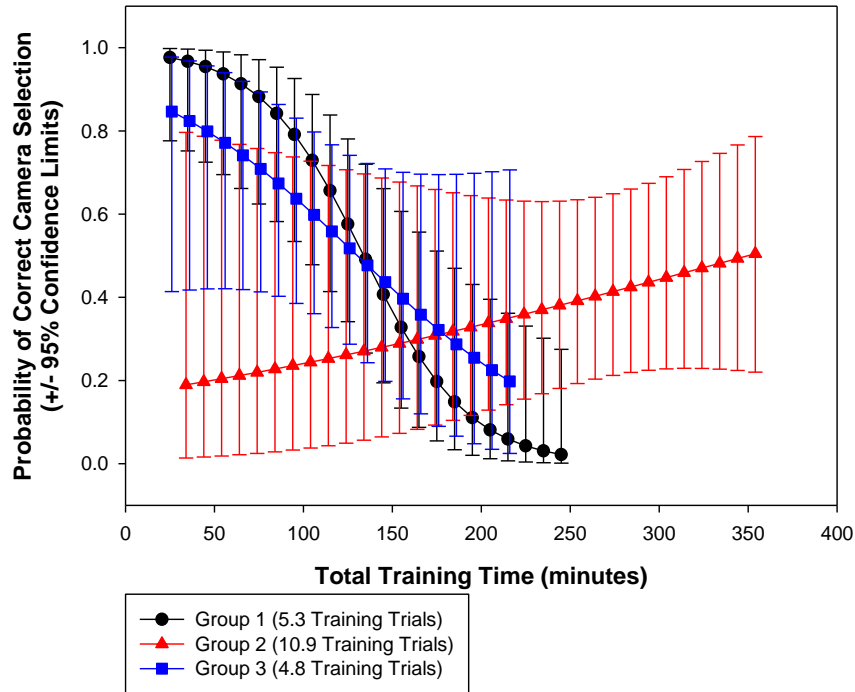
Figure 4.11: Probability of Correct Camera Selection by Group across Total Training Time

The probability of correctly selecting the camera for each level of the trees is summarized in

Table 4.20 along with the odds of correct camera selection versus incorrect camera selection.

The (adjusted) odds ratios comparing the trees are summarized in Table 4.21. The odds of

identifying a target were significantly greater for tree 2 versus trees 1 ($p = 0.0024$) and 4 ($p =$

0.0092) and for tree 3 versus trees 1 ($p < 0.0001$) and 4 ($p < 0.0001$); there was not a significant

difference between trees 2 and 3 ($p = 0.2699$) or between trees 1 and 4 ($p = 0.1423$).

Table 4.20: Probability and Odds of Correct Camera Selection by Group

	Proportion				Odds	
Tree	Estimate	SE	95% CI		Estimate	95% CI
1	0.0723	0.0360	(0.0260, 0.1853)		0.078	(0.027, 0.228)
2	0.9228	0.1047	(0.3838, 0.9957)		11.958	(0.623, 229.6)
3	0.6883	0.0843	(0.5012, 0.8292)		2.209	(1.005, 4.855)
4	0.1610	0.0631	(0.0702, 0.3278)		0.192	(0.076, 0.488)

SE = standard error; CI = confidence interval

Table 4.21: Odds Ratios for Correct Camera Selection Comparing the Trees

Tree	OR	95% CI	p-value
2 vs. 1	153.42	(6.470, 3637.8)	0.0024
2 vs. 3	5.414	(0.258, 113.53)	0.2699
2 vs. 4	62.304	(2.916, 1331.3)	0.0092
3 vs. 1	28.336	(8.821, 91.023)	< 0.0001
3 vs. 4	11.507	(3.782, 35.015)	< 0.0001
4 vs. 1	2.462	(0.731, 8.296)	0.1423

4.7.6 Selection of Port

4.7.6.1 Statistical Method

The probability of correct port selection was modeled at the tree level with a generalized linear mixed effects model (GLMM) assuming a binary distribution for the response (correct port selected) and a logit link function $(X\beta=\ln(\mu/1-\mu))$ that provides the relationship between the linear predictor (model effects)and the mean of the distribution function (selection of port). This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy

91

experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which groups were significantly better for the specific performance measure analyzed..Effects were interpreted using odds ratios describing how many times greater the odds of correct port selection (versus incorrect port selection) was for one group versus another group, where the odds are defined as the probability of correct port selection divided by the probability of incorrect port selection.

4.7.6.2 Results

The GLMM indicated there significant differences in the probability of correct port selection between the groups (p = 0.0244), after controlling for tree (p = 0.0009). There was also no evidence of significant interaction effects for group by tree (p = 0.9592), group by total training time (p = 0.2737), or group by total number of training trials (p = 0.3603), or effects due to previous laparoscopy experience (p = 0.6551), total training time (p = 0.8545), total number of training trials (p = 0.5557), sex (p = 0.0994), or previous gaming experience (p = 0.1647).

The probability of correct port selection for each of the groups is plotted in Figure 4.12 and is summarized in Table 4.22 along with the odds of correct port selection versus incorrect port selection. The odds of correct port selection were 2.651 times greater for group 2 versus group 1 (p = 0.0100) and 2.111 times greater for group 2 versus group 3 (p = 0.0456); the odds were not significantly different between groups 1 and 3 (p = 0.5083).

92
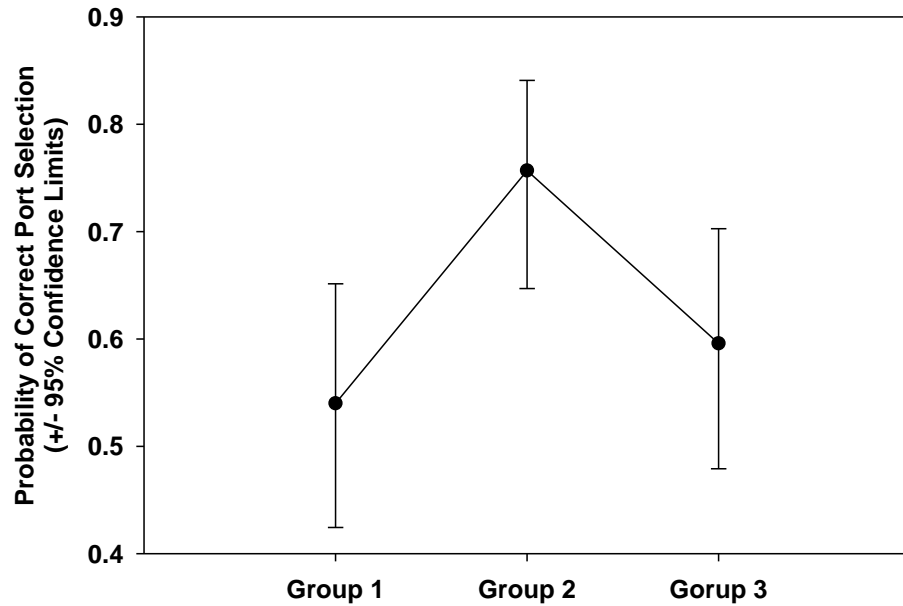
Figure 4.12: Probability of Correct Port Selection by Group

The probability of correct port selection for each level of the trees is summarized in Table 4.23

along with the odds of correct port selection versus incorrect port selection.

Table 4.22: Probability and Odds of Correct Port Selection by Group

	Proportion			Odds	
Tree	Estimate	SE	95% CI	Estimate	95% CI
1	0.8416	0.0546	(0.6998, 0.9237)	5.313	(2.332, 12.105)
2	0.4140	0.0712	(0.2814, 0.5605)	0.707	(0.392, 1.275)
3	0.7157	0.0648	(0.5702, 0.8269)	2.517	(1.326, 4.776)
4	0.4996	0.0726	(0.3575, 0.6416)	0.998	(0.557, 1.791)

SE = standard error; CI = confidence interval

The odds of correct port selection were significantly greater for tree 1 versus trees 2 ($p = 0.0002$) and 4 ($p = 0.0042$) and for tree 3 versus trees 2 ($p = 0.0066$) and 4 ($p = 0.0478$); there

was not a significant difference between trees 1 and 3 ($p = 0.1868$) or between trees 2 and 4 ($p = 0.4185$).

93

Table 4.23: Odds Ratios for Correct Port Selection Comparing the Trees

Tree	OR	95% CI	p-value
1 vs. 2	7.519	(2.757, 20.505)	0.0002
1 vs. 3	2.111	(0.688, 6.481)	0.1868
1 vs. 4	5.322	(1.739, 16.286)	0.0042
3 vs. 2	3.562	(1.449, 8.756)	0.0066
3 vs. 4	2.521	(1.009, 6.299)	0.0478
4 vs. 2	1.413	(0.603, 3.311)	0.4185

4.7.7 Smoothness of Motion

4.7.7.1 Statistical Method

The smoothness of motion was modeled at the tree level with a generalized linear mixed effects model (GLMM) assuming a Poisson distribution for the response (number of major collisions) and a $(X\beta=\ln(\mu/1-\mu))$ that provides the relationship between the linear predictor (model effects)and the mean of the distribution function (responses :smoothness of motion). This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model and later a backward selection process was used to remove any of the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05). If there was evidence of differences between the groups, specific contrasts were tested to determine which groups had significantly different mean smoothness of motion. Effects were interpreted using ratios of the means describing how many times greater the mean smoothness of motion was for one group versus another group.

94

4.7.7.2 Results

The GLMM indicated that the effect of group on the smoothness of motion depends on the total training time ($p = 0.0163$), after controlling for tree ($p = 0.0200$) and total number of training trials ($p < 0.0001$). There was no evidence of significant interaction effects for group by tree ($p = 0.2392$) or total number of training trials by group ($p = 0.8991$), or effects due to previous laparoscopy experience ($p = 0.6024$), previous gaming experience ($p = 0.1280$), port selection ($p = 0.7919$), camera selection ($p = 0.3867$), total training time ($p = 0.6194$), or sex ($p = 0.7371$).

The mean smoothness of motion for each of the groups across the observed values of total training time is plotted in Figure 4.13. Subjects in Group 1, who trained for 23 to 243 minutes, exhibited nominal decreases in the smoothness of motion with increases in training time ($p = 0.6311$), whereas subjects in group 2 ($p = 0.0020$), training for 36 to 361 minutes, and group 3 ($p = 0.0007$), training for 62 to 219 minutes, showed significant decreases in the smoothness of motion with increases in the total training time. Furthermore, increases in the total trials attempted was associated with significant increase in the smoothness of motion ($p < 0.0001$)

The mean smoothness of motion for each of the trees is summarized in Table 4.24. The ratios of the means comparing the trees and previous laparoscopy groups are summarized in Table 4.25. The smoothness of motion was significantly greater for tree 4 versus tree 3 ($p = 0.0053$) and for tree 2 versus tree 3 ($p = 0.0198$); there were not significant differences between trees 1 and 2 ($p = 0.3033$), trees 1 and 3 ($p = 0.3106$), trees 1 versus 4 ($p = 0.1959$), or trees 4 versus 2 ($p = 0.6440$). Furthermore, increases in the total trials attempted was associated with significant increase in the smoothness of motion ($p < 0.0001$)

Table 4.24: Mean Smoothness of Motion by Tree

	Mean		
Tree	Estimate	SE	95% CI
1	77.1304	5.5949	(66.6249, 89.2924)
2	81.8014	6.6717	(69.4217, 96.3887)
3	72.0356	6.2098	(60.5753, 85.6640)
4	83.9823	5.6786	(73.2434, 96.2956)

SE = standard error; CI = confidence interval

Table 4.25: Ratios of Mean Smoothness of Motion Comparing the Trees

Tree	Ratio	95% CI	p-value
1 vs. 2	0.9429	(0.8412, 1.0569)	0.3033
1 vs. 3	1.0707	(0.9363, 1.2245)	0.3106
1 vs. 4	0.9184	(0.8059, 1.0466)	0.1959
4 vs. 2	1.0267	(0.9158, 1.1509)	0.6440
4 vs. 3	1.1658	(1.0495, 1.2951)	0.0053
2 vs. 3	1.1356	(1.0215, 1.2623)	0.0198

CI = confidence interval

4.8 Derived Measures

4.8.1 Statistical methods

All the derived measures path length, collisions and time all per target and collisions per path length, were modeled at the tree level using a GLMM model assuming a Poisson distribution for the responses and log link function ($X\beta = \ln (s)$) that provides the relationship between the linear predictor (model effects)and the mean of the distribution function (responses). This model also took into account the variations in the responses within as well as between subjects for each of the performance measures. The main effects for group and tree and the interaction effects for group by tree, training time by group and number of training attempts by group were included in the model. The fixed effects for previous gaming experience, previous laparoscopy experience, sex, total training time, total number of training trials attempted, camera selected, port selected were initially included in the model. A backward selection process was used to remove any of

96

the effects except that of group and tree if they did not contribute in improving the model fit significantly (p-values ≥ 0.05).If after this process any evidence of differences were observed between groups , specific contrasts were tested to determine which groups were significantly better for the specific performance measure analyzed. Finally effects were interpreted using ratios of the means for each of the measure to determine how many times greater the mean of that measure for one group was compared to the other group

4.8.2 Results

4.8.2.1 Path length per target

The GLMM indicated that the effect of group on the path length per target depends on the total number of training trials (p = 0.0009) and on the total training time (p < 0.0001), after controlling for tree (p < 0.0001) and port selection (p = 0.0029). The model did not find any evidence of significant interaction effects for group by tree (p = 0.1287), or effects due to previous gaming experience (p = 0.5693), previous laparoscopy experience (p = 0.2637), sex (p = 0.8664), or camera selection (p = 0.6437).

Subjects in Group 1exhibited significant decreases in the path length per target with increases in training trial attempts (p < 0.0001) whereas subjects in group 2 and group 3 showed trends for the path length per target to nominally decrease with increases in the number of training trial attempts (Group 2 p = 0.2884 and Group 3 p = 0.4349). The plot for the mean path length per target for each of the groups across the observed values of total number of training trials attempted is shown in (Figure 4.13).
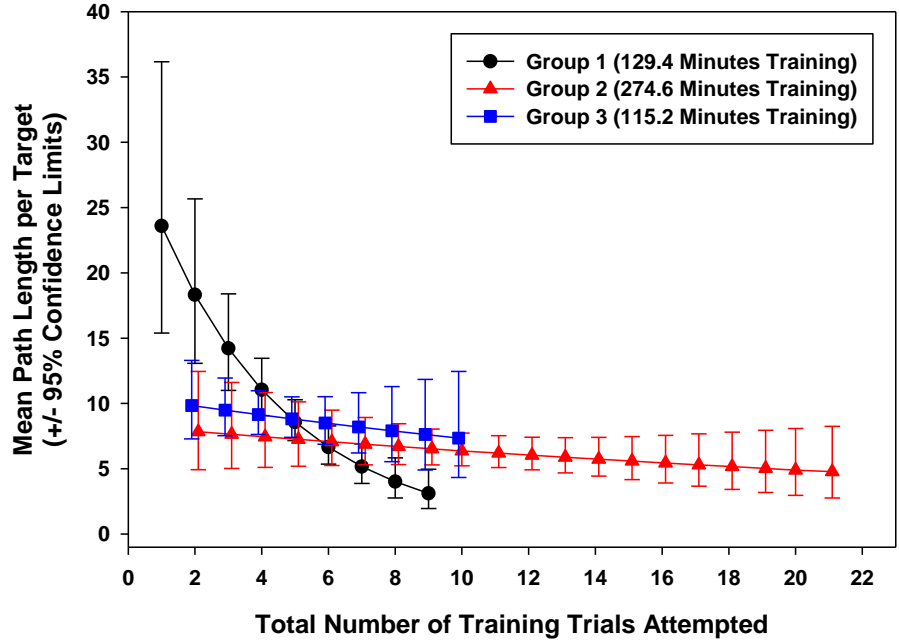
Figure 4.13: Mean Path Length per Target by Group across Total Number of Training

Trials Attempted

In general, subjects in Group 1 exhibited significant increases in the path length per target with

increases in the total training time ($p < 0.0001$) whereas subjects in group 2 and group 3 showed

trends for the path length per target to nominally increase ($p = 0.0741$) and decrease ($p =$

0.6035), respectively, with increases in the total training time. The mean path length per target

for each of the groups across the observed values of total training time is shown in Figure 4.14.
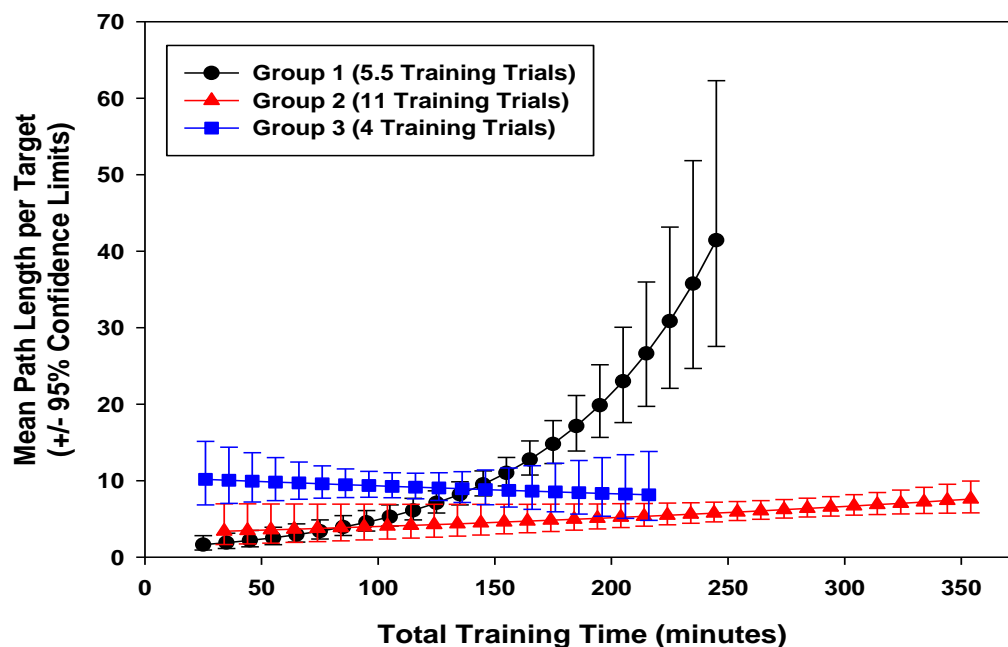
Figure 4.14: Mean Path Length per Target by Group across Total Training Time

The mean path length per target for each level of the trees and port selections is summarized in

Table 4.26. The (adjusted) ratios of the means comparing the trees and previous laparoscopy

groups are summarized in Table 4.27. The mean path length per target was significantly greater

for tree 4 versus trees 1 ($p < 0.0001$) and 2 ($p = 0.0037$) and for tree 3 versus 1 ($p = 0.0089$);

there were not significant differences between trees 3 and 4 ($p = 0.7884$), between trees 2 and 3

($p = 0.0894$), or between trees 1 and 2 ($p = 0.0670$).

Table 4.26: Mean Path Length per Target by Tree and Port Selection

	Mean		
Tree	Estimate	SE	95% CI
1	5.5128	0.4965	(4.6027, 6.6028)
2	6.5800	0.7874	(5.1797, 8.3589)
3	8.8462	1.5044	(6.2898, 12.4415)
4	9.2819	0.8548	(7.7181, 11.1625)

99

Table 4.27: Ratios of Mean Path Length per Target Comparing the Trees

Tree	Ratio	95% CI	p-value
4 vs. 1	1.6837	(1.3941, 2.0335)	< 0.0001
4 vs. 2	1.4106	(1.1245, 1.7695)	0.0037
4 vs. 3	1.0493	(0.7332, 1.5015)	0.7884
3 vs. 1	1.6047	(1.1326, 2.2734)	0.0089
3 vs. 2	1.3444	(0.9538, 1.8949)	0.0894
2 vs. 1	1.1936	(0.9872, 1.4432)	0.0670

The number of major collisions was significantly greater for port selection L as compared to port

selection C (p = 0.0029).

Table 4.28: Mean Path Length per Target by Tree and Port Selection

	Mean		
	Estimate	SE	95% CI
Port Selected			
C	6.4669	0.5955	(5.3780, 7.7764)
L	8.4391	0.8469	(6.9105, 10.3057)

SE = standard error; CI = confidence interval

Table 4.29: Ratios of Mean Path Length per Target Comparing the Trees

Tree	Ratio	95% CI	p-value
Port Selection			
L – C	1.3050	(1.0979, 1.5510)	0.0029

CI = confidence interval

4.8.2.2 Path Length per Collision

The GLMM indicated that there were not significant differences in the path length per collision

between the groups (p = 0.2650), after controlling for tree (p = 0.2649), total number of training

trials (p = 0.0372), and port selection (p = 0.0370). There was also no evidence of significant

interaction effects for group by tree (p = 0.1074), group by total number of training trials (p =

100

0.0845), group by total training time (p = 0.5681), or effects due to previous gaming experience (p = 0.9632), previous laparoscopy experience (p = 0.4708), sex (p = 0.6167), total training time (p = 0.4615), or camera selection (p = 0.9691).

While group 1 nominally had a lowest mean path length per collision and group 3 nominally had the highest, none of the pair wise comparisons were statistically significant (as evident by the test of the group main effect: p = 0.2650). The mean path length per collision for each of the groups is plotted in Figure 4.15 and is summarized in Table 4.30.
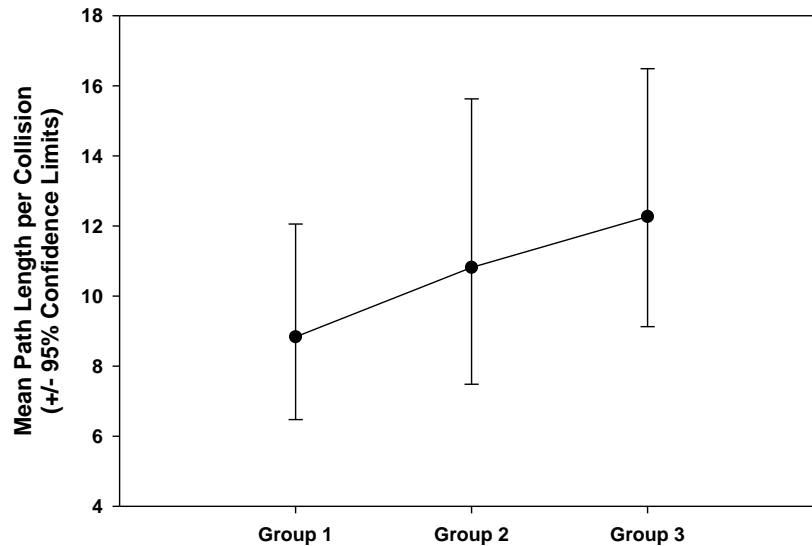
Figure 4.15: Mean Path Length per Collision by Group

Table 4.30: Mean Path Length per Collision by Group

Group	Mean	SE	95% CI
1	8.8339	1.3241	(6.4744, 12.0534)
2	10.8142	1.9775	(7.4840, 15.6264)
3	12.2644	1.7891	(9.1244, 16.4850)

CI = confidence interval

101

There was a significant effect of number of training trials on the mean path length per collision (p = 0.0372). In general, increases in the number of training trials were associated with increases in the mean path length per collision.

There was also a significant effect of port selection on the mean path length per collision (p = 0.0370); the path length per collision was 1.3272 times greater for port L as compared to port C (Tables 4.31).

Table 4.31: Ratios of Mean Path Length per Collision Comparing the Port Selection Groups

Port Selection	Ratio	95% CI	p-value
L – C	1.3272	(1.0179, 1.7306)	0.0370

CI = confidence interval

4.8.2.3 Collisions per Target

The GLMM indicated that the effect of group on the collisions per target depends on the total number of training trials (p = 0.0095), after controlling for tree (p = 0.0447). There was no evidence of significant interaction effects for group by tree (p = 0.6942), group by total training time (p = 0.7086), or effects due to total training time (p = 0.8003), previous gaming experience (p = 0.7845), previous laparoscopy experience (p = 0.0725), sex (p = 0.1513), camera selection (p = 0.1685), or port selection (p = 0.6053).

In general, subjects in Group 1 exhibited significant increases in the collisions per target with increases in training trial attempts (p = 0.0048) whereas subjects in group 2 (p = 0.6370) and group 3 (p = 0.1767) showed trends for the mean collisions per target to nominally decrease with increases in the number of training trial attempts. The mean collision per target for each of the

102

groups across the observed values of total number of training trials attempted is shown in Figure 4.16.
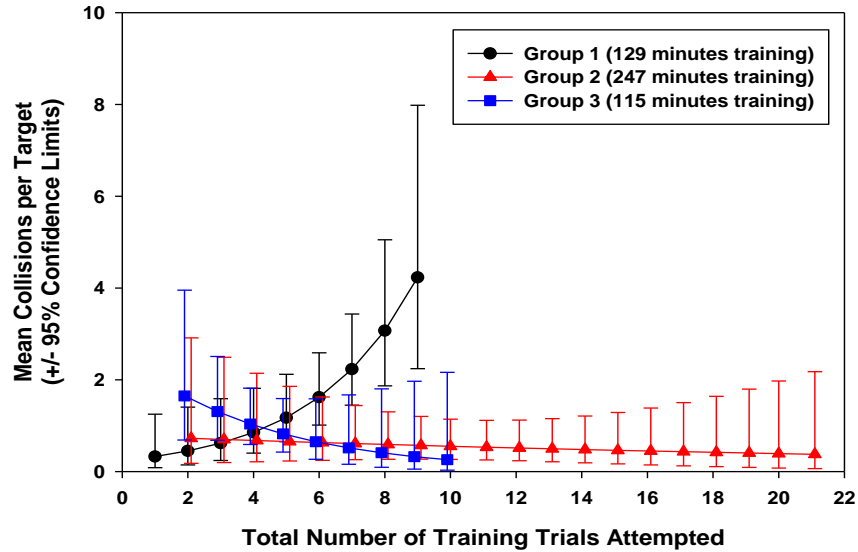

Figure 4.16: Mean Collision per Target by Group across Total Number of Training Trials Attempted

The mean collision per path length for each of the trees is summarized in Table 4.32. The (adjusted) ratios of the means comparing the trees are summarized in Table 4.33. The mean collision per path length was significantly greater for tree 1 versus trees 2 ($p = 0.0094$) and 3 ($p = 0.0011$) and for tree 4 versus 3 ($p = 0.0199$); there were not significant differences between trees 1 and 4 ($p = 0.2559$), between trees 2 and 4 ($p = 0.0557$), or between trees 2 and 3 ($p = 0.2622$). There was also evidence that total training time has a significant effect on the mean collisions per path length ($p = 0.0411$). In general, increases in training time was associated with decreases in the mean collisions per path length

Table 4.32: Mean Collisions per Path Length by Tree

Tree	Mean		
	Estimate	SE	95% CI
1	0.1655	0.0459	(0.0949, 0.2884)
2	0.0866	0.0160	(0.0597, 0.1257)
3	0.0683	0.1987	(0.0381, 0.1222)
4	0.1259	0.0272	(0.0817, 0.1939)

SE = standard error; CI = confidence interval

Table 4.33: Ratios of Mean Collisions per Path Length Comparing the Trees

Tree	Ratio	95% CI	p-value
1 vs. 2	1.9103	(1.1811, 3.0899)	0.0094
1 vs. 3	2.4243	(1.4505, 4.0518)	0.0011
1 vs. 4	1.3147	(0.8147, 2.1216)	0.2559
4 vs. 2	1.4530	(0.9905, 2.1315)	0.0557
4 vs. 3	1.8440	(1.1067, 3.0724)	0.0199
2 vs. 3	1.2690	(0.8318, 1.9361)	0.2622

CI = confidence interval

4.8.2.4 Time per Target

The GLMM indicated that the effect of group on the time per target depends on the total number of training trials ($p = 0.0067$) and the total training time ($p = 0.0116$), after controlling for tree ($p = 0.0007$). There was no evidence of significant interaction effects for group by tree ($p = 0.3227$), or effects due to previous gaming experience ($p = 0.9472$), previous laparoscopy experience ($p = 0.8982$), sex ($p = 0.1987$), camera selection ($p = 0.7976$), or port selection ($p = 0.2591$).

In general, subjects in Group 1 ($p < 0.0001$) and Group 3 ($p = 0.0172$) exhibited significant decreases in the time per target with increases in training trial attempts ($p < 0.0001$) whereas

104

subjects in group 2 showed trends for the time per target to nominally decrease with increases in the number of training trial attempts (p = 0.0756). The mean time per target for each of the groups across the observed values of total number of training trials attempted is shown below in Figure 4.17.
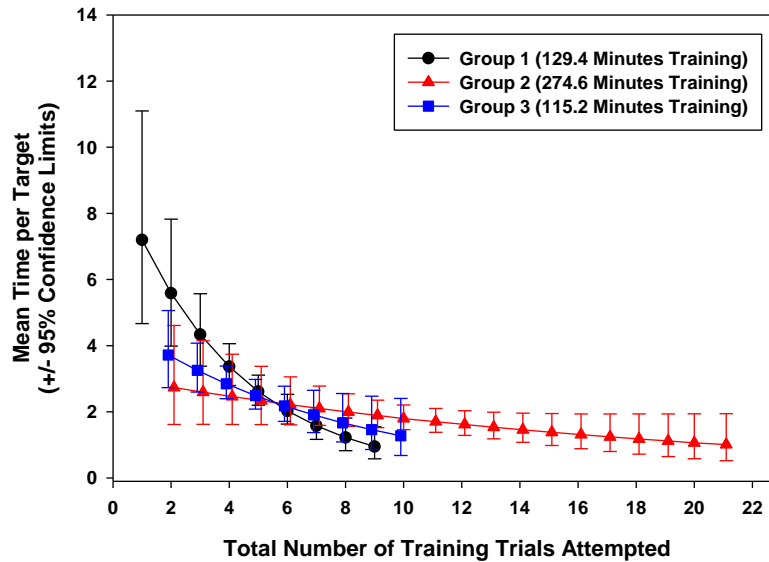
Figure 4.17: Mean Time per Target by Group across Total Number of Training Trials Attempted

In general, subjects in Group 1 (p < 0.0001) and Group 2 (p = 0.0242) exhibited significant increases in the time per target with increases in the total training time whereas subjects in group 3 showed trends for the time per target to nominally increase (p = 0.1266) with increases in the total training time. The mean time per target for each of the groups across the observed values of total time training is shown below in Figure 4.18.

105
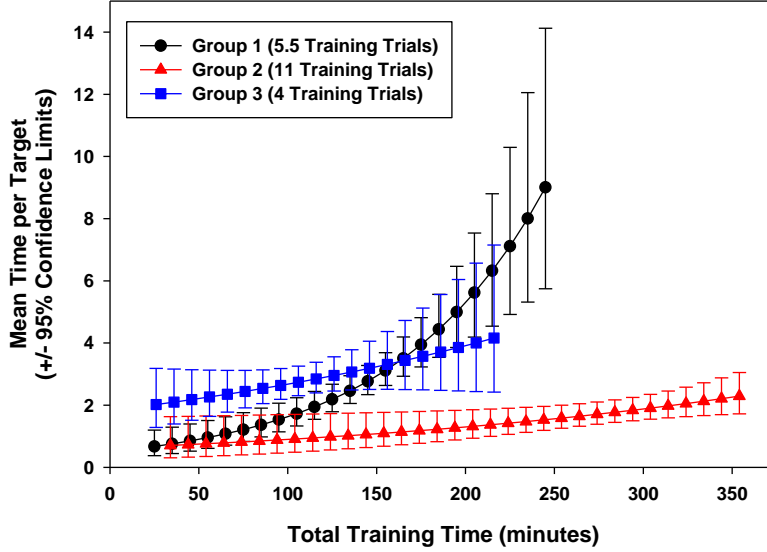
Figure 4.18: Mean Time per Target by Group across Total Training Time

4.8.2.5 Cognitive Load Analysis

The last variable that we used in our analysis was the cognitive load of the participants during the testing task for all the three groups. A linear mixed effects model was used to test for differences in the cognitive load response (NASA TLX) between the groups, adjusting for tree. There was no evidence of a significant interaction between group and tree (p = 0.8707). There was no evidence of significant differences in cognitive load among the groups (p = 0.2743), after controlling for tree (p = 0.0028). The mean cognitive load is summarized in Table 4.34 and Figure 4.19 for each of the groups.

Table 4.34: Mean Cognitive Load by Group

Group	Mean	SE	95% CI
1	53.9984	5.7059	(41.3064, 66.6903)
2	50.5291	3,9232	(41.7359, 59.3224)
3	60.1337	3.5859	(52.5000, 67.7674)

CI = confidence interval

106
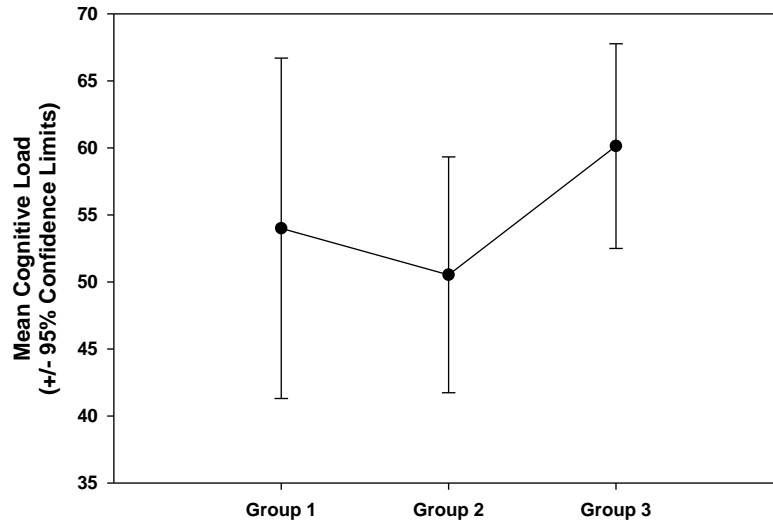
Figure 4.19: Mean Cognitive Load by Group

Table 4.35: Differences in Mean Cognitive Load Comparing the Groups

Comparison	Difference	SE	95% CI	p-value
Group 1 – 2	3.4692	6.8370	(-10.9693, 17.9077)	0.6185
Group 3 – 1	6.1353	6.7674	(-8.1748, 20.4454)	0.3777
Group 3 – 2	9.6046	5.7056	(-2.2631, 21.4722)	0.1072

CI = confidence interval

107

Chapter 5

Discussion

The primary objective of this study was to evaluate two different training methods in terms of their effectiveness in helping learners perform a laparoscopic camera navigation task in a stressful environment. We compared three training methods: (1) training under no stress until high performance levels are reached (the control condition), (2) training under no stress until high performance levels and low cognitive load are reached and (3) training under a different stressor than the one in the test environment until high performance levels are reached.

The primary performance variables determined for the testing task were: the number of targets identified, the number of major collisions committed, time taken to complete the task, path length taken to find the targets, the selection of camera and port and smoothness of motion. These were analyzed as they are important factors to assess whether a training method was effective. Similarly, derived measures such as path length per target, path length per collision, collision per path length, collision per target, time per target were also used in the analysis. In

108

addition to the performance variables, the cognitive load for each of the subjects was also recorded after each trial during the testing task using the NASA TLX.

5.1 Proportion of Targets Identified

The results of the study confirm that subjects who were trained to low cognitive load for the training task performed better when subjected to perform similar tasks in stressful conditions. Comparing the three groups for the number of targets indentified, the Group 2 subjects found 1.65 and 1.53 times the number of targets identified by those in Group 1 and Group 3, respectively. There was also no significant difference between Group 1 and Group 3 for this performance measure, indicating that training under stress did not help in improving performance for this variable.

The statistical model also revealed that the better performance was not achieved due to the longer training time of G2, which was 2.1-2.4 times greater than G1 and G3 respectively (G2:274 min vs.G1:129min and G3:115min), and higher number of training trials attempted, which was 2-2.75 times greater for G2 than those in G1 and G3 (G2:11 vs. G1:5.5 and G3:4): there was no evidence of effects due to total training time, the number of training trials and interaction effects with group. This confirms that the design of the training method resulted in better performances during the testing task rather than these other variables.

In addition, as expected, those with laparoscopic experience performed better than those without indicating that experts were better at the laparoscopic task than novices. This is consistent with the literature. One important result that we did not find to be consistent with other literature findings was that of the effect of gaming experience on performance in that we found no significant difference. This may be that more accurate consideration of the amount of

109

time of the gaming experience and the games played may be needed. We also observed that subjects were able to identify targets 1.685 times greater for 30 degree laparoscope compared to 0 degree laparoscope. This was also surprising as typically, when training, the 30 degree laparoscope is found to be difficult to use.

5.1.1 Effect of target difficulty and tree difficulty on proportion of targets identified

The proportion of targets identified in each of three groups (easy, medium, hard) was directly proportional to the degree of difficulty as expected: a higher number of easy targets were identified compared to medium and hard targets (Easy: 0.66 vs. Medium: 0.45 vs. Hard: 0.51). However, although all the trees were designed to be equally difficult, there were significant differences in the proportion of targets identified based on which tree was examined. Subjects were able to identify the most number of targets for Tree 2 and the least number of targets for Tree 4. It is unclear what produced the inconsistency between target and tree difficulty. However, one overall effect that would affect performance with a tree is whether the participant correctly selected the camera and port for a specific tree. If not, the tree would seem harder than it really was.

5.2 Cognitive Skill Development

One of the goals of the study was to determine if requiring participants to use different port locations and camera angle would enable the learners to develop the underlying conceptual relationship between them and the position of targets within the trainer box. This would mean that they would then be able to choose the correct port and camera angle for an actual surgery when presented with a set of images (i.e., CT scan, MRI, etc.) as in the operating room. We determined the effectiveness of the different training methods on the cognitive skill level attained

by comparing the percentage of correctly selected camera angles and port selections for each of the testing task. As described above, his skill also is expected to affect the proportion of targets identified.

5.2.1 Camera selection

While there were no significant differences between the three training methods, there existed a strong positive relationship between the number of training trials and percentage correct selection of the camera angle in the testing task for Group 1, and a strongly negative relationship between total training time and percentage correct. This would suggest that those participants that trained with many short trials were better at performing the camera selection. This may suggest a training paradigm which could be examined in the future to see if this strategy produced better performance. The same trends did not exist for Group 2 and Group 3 and, in fact, were not statistically significant. This suggests that for these training methods, the training paradigm of short versus long trials does not have an effect. In some sense this is a positive result as it suggests that the stopping criteria, as given, are adequate to bring learners of diverse ability to the same level of performance.

5.2.2 Port selection

A more definitive relationship seems to exist between the training methods and correct selection of a port. The effect of training method on the percentage correct selection of port revealed that subjects in Group 2 performed better than the other two groups. Comparison between the groups revealed that participants in Group 2 were able to select the correct port location 2.651 times greater than subjects in Group 1 and 2.11 times greater than subjects in Group 3. No significant difference was observed for this variable between Groups 1 and 3. These results support the

111

assertion that training until automation enables learners to retain the higher level cognitive skill required in making decisions under stress better than the other training methods. However, this result is weakened by the observation that there was not an effect between groups for camera selection, and average performance was relatively low.

5.2.3 Effect of different tree setups on camera angle and port selection

We observed that both correctly selecting the camera angle and port location varied significantly with the different tree setups. We observed that the probability of correctly selecting the camera angle was highest for Tree 2 and lowest for Tree 1.Conversely the subjects were able to correctly identify the port location the most number of times for Tree 1 and the lowest number of times for Tree 2.

5.3Effect of the different training approaches on the participants' motor skills

To examine the effectiveness of the three training methods on a participant's motor skills while performing under stress, the time taken to complete the task, the mean path length and the smoothness of motion were analyzed.

5.3.1 Comparison of total time taken by groups

Overall, subjects in Group 2 took less time to find the targets compared to those in Group 1 and Group 3 (6.4 vs. 9.0 vs. 9.1) but the differences were not statistically significant. This supports the fact that the number of targets found by Group 2 was not simply due to longer trial times for the testing task, as the reverse trend would be expected. Also note that, even in the ideal case where a participant moves straight from target to target, a finite amount of time is needed for

each target and, therefore, we would expect actually a longer total time. The fact that it is shorter, suggests better performance with Group 2.

We also observed a positive effect of the number of training trials on this variable across the groups, where subjects who attempted a higher number of trials required less time to complete the task during the testing. This, again, suggests that a possible training paradigm is to encourage learners to attempt a larger number of short trials, rather than persist for a long time on any given trial. In contrast, a negative relationship seems to exist between the total training time and time to complete the testing task.

5.3.2 Comparison of path length for the three groups

The results for the path length variable indicate that group 2 took a shorter mean path length (30% shorter) as compared to Group 1 and group 3. However, as the differences are not statistically significant, at first glance, the effect of training seems to be neutral for this variable. If one considers, though, that significantly more targets were found for comparable path lengths, this suggests an improvement in motor performance for Group 2. Note that to reach any target takes a finite path length, so even in the optimal case, path lengths would be longer for more targets found.

However, the training time had a negative effect on mean path length, with longer training times associated with longer path lengths. This may be because the poorer learner had a longer training time which is reflected in the mean path length. This raises caution in terms of backing out a training paradigm based on the higher performing learners. It may be that the parameter is correlated but not causal of the results. This shows that further evaluation is really necessary to show that training participants to have many short trials would be of benefit.

113

Overall there was a positive effect of training on the targets identified while the effect is neutral with respect to time taken and mean path length.

5.3.2.1 Effect of laparoscopic experience on mean path length

From the results it is evident that the mean path length was 1.522 times higher for those with previous laparoscopic experience compared to those with no experience (Yes: 34.17 vs. No: 22.15). This was surprising as we expected the reverse trend. However, those with previous laparoscopic experience found more targets, each of which, in the optimal condition, still requires a finite movement distance to complete.

5.3.3 Comparison of Jerk a measure of Motion smoothness for the different groups

Another motion metric that we analyzed was the smoothness of motion, which is an important metric that has been used to differentiate the expert surgeons from novices (Stylopoulos et al 2004). The results of this study indicated that the smoothness of motion varied across training time for each of the 3 groups. Training time had a positive effect on performance for all here groups, with Groups 2 and 3 having significant improvements in this variable compared to the nominal improvements for subjects in Group 1. The smoothness of motion decreased with increases in training trials for all the groups.

It should be noted that the smoothness of motion was better for Groups 1 and 3, as compared to Group 2 for the same number of training trials. This is contradictory to our expectations. However, other work in our laboratory has suggested that smoothness of motion may decrease for better performance due to the increased speed of the movements.

114

5.3.3.1 Effect of testing tree on smoothness of motion

We observed variations in the smoothness of motion variable for the different trees. The subjects performed with higher level of smoothness when attempting Tree 4 and Tree 2 and less smoothly when attempting Tree 3(T4:83.98 vs. T2:81.80 vs. T1:77.13 vs. T3:72.03)

5.4 Effect of the training approaches on Path length per target and Time per target (a measure of speed)

Examination of the derived variable path length per target revealed a complex relationship with the number of training trials and total training times. Longer total training times seem to have a negative effect on performance while greater number of training trials causes improvement in performance. These effects were pronounced and statistically significant for subjects in Group 1 in contrast to the other two groups. This is a positive result for Groups 2 and 3 as it suggests that the resulting performance during the testing task is relatively independent of total training time and number of trials, as opposed to Group 1. For actual training, we would like to see learners achieve the same performance regardless of these two measures, as they are not the bases for stopping training.

Examining the effect of training time and trials attempted for the time per target variable indicated that the longer the training time the slower the identification of a target, whereas the greater the number of training trials attempted the faster the identification of the target. These effects were pronounced for subjects in Group 1 in contrast to the other two groups. The results reveal that there exists a complicated relationship between this variables with respect to the training characteristics and requires further study to better understand the underlying relationship. However, it does suggest, as with path length per target, that training participants to

115

use a paradigm where they explore multiple trees quickly in succession, rather than fewer trees more thoroughly might be beneficial.

5.5 Comparison of the number of major collision by groups

One very important requirement while performing laparoscopic surgeries is to avoid committing errors, such as damaging collisions with soft tissue, as this can adversely impact the outcome of the surgery. Hence effective training needs to ensure the minimization of any damage while performing under stressful conditions .This was assessed by comparing the three groups on the number of major collisions with the task trees during the testing task. There was no simple outcome for this variable. The mean number of collisions was found to nominally decrease with increasing number of training trials for subjects in Group 2 and 3, whereas it was found to significantly increase with training trials for subjects in Group 1. It should be noted that the trends in Group 2 and 3 were only nominal, with both producing a relatively constant (and similarly low) value. This suggests that these two methods are a good training design as training produces the same performance regardless how people train (i.e., training does succeed in having a leveling effect) and the performance result is good.

One reason for the effect as a function of number of training trials could be that the stress in the testing condition could have impacted performance, at least for this measure for Group 1. This could be because subjects who had a longer learning curve for this performance measure could have required additional mental resources to handle the stress component in testing, which could have, subsequently, resulted in degradation of testing performance for this variable. Further examining the collision per target variable reveals that performance of subjects in Group 1 is negatively affected by increases in the number of training trials while it is positive for the

116

other two groups. This could once again be due to the effect of stress on performance which is one of the main reasons that alternate training methods need to be developed to handle the effect of stress.

5.5.1 Effect of tree on the number of major collisions

We observed the number of major collisions to vary with the testing tree with highest number of collisions when attempting Tree 4 and lowest in Tree 3 (T1:3.61 vs. T2:2.45 vs. T3:1.73 vs. T4:4.0).

5.5.2 Effect of laparoscopic experience on the number of major collisions

Subjects with previous laparoscopic experience committed 2.5 times more number of collisions compared to those without (Yes: 6.77 vs. No: 2.7). This was surprising as we would have expected these individuals to perform fewer errors. This may have been due to their awareness that there was no real consequence of their having collisions in terms of endangering a patient.

5.5.3 Comparison to collisions per path length by groups

Training also seemed to have a negative impact on collisions per path length measures for subjects in Group 1 but not in case of Group 2 and Group 3

5.5.4 Effect of tree on collisions per target and collisions per path length

We observed that the collision per target measure varied with the testing trees and was highest for Tree 4 and lowest for Tree 2(T1:1.02 vs. T2:0.62 vs. T3:0.82 vs. T4:1.18).The collision per path length measure was highest for Tree 1 and lowest for Tree3 (T1:0.16 vs. T2:0.08 vs. T3:0.06 vs. T4:0.12).

117

5.6 The effect of training and of stress on the cognitive load of a subject

Finally, we expected the cognitive load scores to be low for the subjects in Group 2 during testing as compared to Groups 1 and 2. However, the results from the statistical model concluded that although the mean value of the NASA TLX scores were nominally lower for Group 2 as compared to the other two groups, they were not statistically significant. This was not expected as the primary objective of training until low cognitive load was to ensure that the skills learnt get automated, ensuring the available mental resources to handle the stress component. Although the stress component would be added to the cognitive load, one would expect it to have an effect on, at least, participants in Group 1, who had no stress training, as well. It was also surprising that participants in Group 3 did not have a significantly lower cognitive load as well, as they would, if training for stress transferred to the novel stressor, have a lower load for the stress factor. One possible reason that scores were not significantly lower for subjects in Groups 2 and 3 is that the NASA TLX score was not sensitive enough to assess small changes in cognitive loading and that other potentially effective metrics of assessment could provide more insight into this.

This study attempted to design and test the efficacy of three training approaches, namely, training until automaticity, stress exposure training and the normal training method, for performing a laparoscopic camera navigation task under stressful conditions. The results from the study indicate that those trained in Group 2 (training to automaticity) have performed better compared to the other groups under stress. From the results we were able to confirm that at least for the most relevant measure, proportion of targets identified, the benefits of the training model was clearly evident since this measure was not affected by the training duration. The total time taken and path length were only nominally better for those in Group 2 as compared to the other

118

two groups. However, together with the fact that more targets were found and a finite time and distance is needed to find each target, also supports that those who trained for automaticity performed better in the testing task. In addition, in the OR, it is more important to maintain accuracy than to complete the task quicker that determines the success of the surgery. Participants in Group 2 also performed noticeably better for port selection (a cognitive task) than those in the other two groups.

We found that for the total time taken and total path length, performance on the testing task improves with the number of training trials performed and decreases with the total training time for all groups. This was also true for camera angle for Group 1 (normal/control condition). These results suggest that a further paradigm might be to encourage learner to perform many short training trials rather than languish in finding targets with longer training trials. As this is not necessarily a causal relationship, further testing is needed to examine whether this can improve performance for all types of learners. However, one measurement, smoothness of motion counters this trend in that an increase in total training time and a decrease in the total number of trials improve the smoothness of motion (the exact opposite of the above). Also, Group 2 is nominal worse than Groups 1 and 3, rather than nominal better than Groups 1 and 3 for total time taken and total path length. Current work in our lab suggests that smoothness of motion may not be an effective variable as Jerk appears to increase with the speed used, which can be considered a sign of better performance.

Finally, for the number of collisions, we obtained a more complex response. The performance of all groups on this measure was a function of the number of training trials performed. With Group 1, a decrease in performance was found for an increase in the number of training trials, whereas for Groups 2 and 3, a nominal increase in performance was found. The

119

result for Group 1 suggested that increasing the number of training trials, by itself, will actually decrease performance, and this argues for the need for different training methods such as methods 2 and 3. The fact that the result was relatively constant (statistically speaking) for Groups 2 and 3, suggest that the stopping criteria were effective for these methods, even though no specific time duration was given.

Chapter 6

Conclusion

While stress is considered one of the major issues that affect performance of a surgeon in the OR, current minimally invasive surgical training simulators have not included in their training a provision for learning how to handle stress. In this thesis we proposed alternative approaches to training in a task that will lead to better performance while performing under stress. We designed two training approaches to handling stress, namely, training until the task is automated and training under stress. The main experiment compared these approaches to the normal training method to determine the better training system of the three.

In order to test these methods, we designed a trainer system that consisted of a simulated laparoscope, a video monitor and various sensors to record and assess participants' performance. We also created a controlled environment where stressors were induced artificially. To ensure that the training provided did what we expected, we ensured that the components of the trainer system were standardized and the stressors used for training and testing were effective. Chapter 2

described the various standardization procedures that were carried out on the components of the trainer system. In order to create a stressful environment while performing the task we did several studies, as discussed in Chapter 3, to determine consistent stressors that can be used in our environment. The various stressors and stress measurement techniques were tested and two stressors, namely, arm stretch and neuromuscular electrical stimulator were identified for this study.

Finally the results from the main experiment described in the previous chapter clearly show that training until automation is the most effective approach to handling stress compared to the normal approach currently in practice. Those trained with this method were found to have better retention skills evaluated in terms of their ability to find more targets and maintain accuracy measured by the number of collisions while doing so, due to their higher resistance to stress effects compared to the other groups. While the results from the study revealed that training until automation did not significantly improve the individual's ability to execute the task faster, the benefits of the surgeon's skill to maintain accuracy in the OR is more important in terms of patient safety as compared to their speed. The effects of training under stress were also found to be ineffective to prepare the subjects to handle a novel stressor. While there were significant differences in the different performance measures between Groups 2 and 3similar differences were not observed between Groups 1 and 3 confirming that training under stress is just as effective as the control group. However similar to those trained until automation (Group 2) training under stress seem to result in the subjects to become resistant to stress effects which seems to be the only benefit of training under stress.

In considering the ability to implement the best method, the use of cognitive load measures to measure the level of automation can easily be incorporated into current trainer

systems, as the use of NASA TLX for cognitive assessment is easy to administer, reliable and fast. Overall this study provides the ground work for future work on developing training methods that can provide a learner a better and more efficient way to learn and perform complicated MIS tasks.

Bibliography

Dimitrios, Stefanidis, MD, PhD, James, R., Korndorffer Jr, MD, FACS, Sarah, Markley, BS†, Rafael, Sierra, MD†, Daniel, J., Scott, MD, FACS‡,. (2006). Proficiency maintenance: Impact of ongoing simulator training on laparoscopic skill retention.202 (4), 599-603.

I., R., Crothers , A., G., Gallagher , N., McClure , D., T., D., James , J., McGuigan . (1999). Experienced laparoscopic surgeons are automated to the "Fulcrum effect": An ergonomic demonstration. Endoscopy, (31), 365-369.

James, R., Korndorffer, Jr, MD (FACS), J., Bruce, Dunne, PhD, Rafael, Sierra, MD, Dimitris, Stefanidis, MD, PhD, Cheri, L., Touchard, BS, Daniel, J., Scott, MD. (2005). Simulator training for laparoscopic suturing using performance goals translates to the operating room .201(1), 23-29.

Natalie, Perkins, Janet, L, Starkes, Timothy, D., Lee, Carol, Hutchison. (2002). Learning to use minimal access surgical instruments and 2-dimensional remote visual feedback: How

difficult is the task for novices?. Advances in Health Sciences Education Theory and Practice, 7(2), 117-118-131.

Ahlberg, G., Enochsson, L., Gallagher, A. G., Hedman, L., Hogman, C., McClusky, D. A.,3rd, et al. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies American Journal of Surgery, 193(6), 797-804.

Andreatta, P. B., Woodrum, D. T., Birkmeyer, J. D., Yellamanchilli, R. K., Doherty, G. M., Gauger, P. G., et al. (2006). Laparoscopic skills are improved with LapMentor training: Results of a randomized, double-blinded study Annals of Surgery, 243(6), 854-60; discussion 860-3.

Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., & Kneebone, R. (2010). The impact of stress on surgical performance: A systematic review of the literature Surgery, 147(3), 318-30, 330.e1-6.

Bacon, S. J. (1974). Arousal and the range of cue utilization. Journal of Experimental Psychology, 102(1), 81 -87.

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development Cognition, 10(1-3), 17 -23.

Bainbridge, L. (1997). The change in concepts needed to account for human behavior in complex dynamic tasks IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 27(3), 351 -359.

125

Belkic, K., & Nedic, O. (2007). Workplace stressors and lifestyle-related cancer risk factors

    among female physicians: Assessment using the occupational stress index Journal of

    Occupational Health, 49(1), 61 -71.

Berguer, R., Smith, W. D., & Chung, Y. H. (2001). Performing laparoscopic surgery is

    significantly more stressful for the surgeon than open surgery Surgical Endoscopy, 15(10),

    1204-1207.

Bohm, B. (2001). A prospective randomized trial on heart rate variability of the surgical team

    during laparoscopic and conventional sigmoid resection Archives of Surgery, 136(3), 305 -

    310.

BOND, A., & LADER, M. (1974; 2011). The use of analogue scales in rating subjective feelings

    British Journal of Medical Psychology, 47(3), 211 -218.

Carney, R. M., Saunders, R. D., Freedland, K. E., Stein, P., Rich, M. W., & Jaffe, A. S. (1995).

    Association of depression with reduced heart rate variability in coronary artery disease The

    American Journal of Cardiology, 76(8), 562-564.

Cascio, W. F., Honig, P., & Deventer, W. v. (1991). Applied psychology in personnel

    management . Englewood Cliffs, N.J.: Prentice Hall.

Chmarra, M. K., Bakker, N. H., Grimbergen, C. A., & Dankelman, J. (2006). TrEndo, a device

    for tracking minimally invasive surgical instruments in training setups Sensors and

    Actuators A: Physical, 126(2), 328-334.

126

Collins, D. L., Baum, A., & Singer, J. E. (1983). Coping with chronic stress at three mile island: Psychological and biochemical evidence. Health Psychology, 2(2), 149 -166.

Combs, A. W., & Taylor, C. (1952). The effect of the perception of mild degrees of threat on performance. The Journal of Abnormal and Social Psychology, 47(2, Suppl), 420 -424.

Cox, D., Hallam, R., O'Connor, K., & Rachman, S. (1983; 2011). An experimental analysis of fearlessness and courage British Journal of Psychology, 74(1), 107 -117.

Crocker, J., Alloy, L. B., & Kayne, N. T. (1988). Attributional style, depression, and perceptions of consensus for events. Journal of Personality and Social Psychology, 54(5), 840 -846.

Dankelman, J. (2008). Surgical simulator design and development World Journal of Surgery, 32(2), 149-155.

Darzi, A., Munz, Y., Dosis, A., Bann, S., & Moorthy, K. (2003). The effect of stress-inducing conditions on the performance of a laparoscopic task Surgical Endoscopy, 17(9), 1481-1484.

Datta, V., Chang, A., Mackay, S., & Darzi, A. (2002). The relationship between motion analysis and surgical technical assessments The American Journal of Surgery, 184(1), 70-73.

Datta, V., Mackay, S., Mandalia, M., & Darzi, A. (2001). The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model1 1No competing interests declared. Journal of the American College of Surgeons, 193(5), 479-485.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women International Journal of Psychophysiology, 37(2), 121 -133.

Dobkin, P. L., & Pihl, R. O. (1992). Measurement of psychological and heart rate reactivity to stress in the real world Psychotherapy and Psychosomatics, 58(3-4), 208 -214.

Driskell, J. E., Hughes, S.,C.,Hall,J.,K., & Salas, E. (1992). Guidelines for implementing stress training. Orlando, Florida: Naval Air Warfare Center Training Systems Division.

Driskell, J. E., & Salas, E. (1991). Group decision making under stress. Journal of Applied Psychology, 76(3), 473 -478.

Driskell, J. E., Salas, E., & Johnston, J. (1999). Does stress lead to a loss of team perspective? Group Dynamics: Theory, Research, and Practice, 3(4), 291 -302.

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. Journal of Applied Psychology, 77(5), 615 -622.

Driskell., J., E., & Salas., E. (1991). Group decision making under stress. Journal of Applied Psychology, 76, 473-478.

Driskell., J., E., Willis., R., & Copper., C. (1992). Effect of overlearning on retention. Journal of Applied Psychology, 77, 615-622.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. Psychological Review, 66(3), 183 -201.

Evans, G, W, & Cohen, S.,. (1987). Environmental stress. In S. D, & A. I (Eds.), <u>Handbook of environmental psychology, vol. I</u> (pp. 571-610). New York: Wiley.

Eyal, R., & Tendick, F. (2001). Spatial ability and learning the use of an angled laparoscope in a virtual environment Studies in Health Technology and Informatics, 81, 146-152.

Fitts,P,,M. (1964). Perceptual motor skills learning. In W. A. Melton (Ed.), Categories of human learning (pp. 243-285). New York: Academic Press.

Fitts, P. M., Weinstein, M., Rappaport, M., Anderson, N., & Leonard, J. A. (1956). Stimulus correlates of visual pattern recognition: A probability approach. Journal of Experimental Psychology, 51(1), 1 -11.

Forsman, L., & Lindblad, L. E. (1983). Effect of mental stress on baroreceptor-mediated changes in blood pressure and heart rate and on plasma catecholamines and subjective responses in healthy men and women Psychosomatic Medicine, 45(5), 435-445.

Fox, M. L., Dwyer, D. J., & Ganster, D. C. (1993). Effects of stressful job demands and control on physiological and attitudinal outcomes in a hospital setting Academy of Management Journal.Academy of Management, 36(2), 289-318.

Fried, G. M. (1999). Comparison of laparoscopic performance in vivo with performance measured in a laparoscopic simulator Surgical Endoscopy, 13(11), 1077-1081.

Haluck, R., Webster, R., Snyder, A., Melkonian, M., Mohler, B., Dise, M., et al. (2001). A virtual reality surgical trainer for navigation in laparoscopic surgery.. Medicine Meets Virtual Reality, 81, 171-6.

Hamilton, E. C., Scott, D. J., Kapoor, A., Nwariaku, F., Bergen, P. C., Rege, R. V., et al. (2001). Improving operative performance using a laparoscopic hernia simulator The American Journal of Surgery, 182(6), 725-728.

Hamstra, S. J., & Dubrowski, A. (2005). Effective training and assessment of surgical skills, and the correlates of performance Surgical Innovation, 12(1), 71-77.

Hanna, G. B., Frank, T. G., & Cuschieri, A. (1997). Objective assessment of endoscopic knot quality The American Journal of Surgery, 174(4), 410-413.

Hans, Z. (1994). Workload and psychophysiological stress reactions in air traffic controllers. Ergonomics, 37(3), 525-539.

Harnod, T., Yang, C. C., Hsin, Y. L., Shieh, K. R., Wang, P. J., & Kuo, T. B. (2008). Heart rate variability in children with refractory generalized epilepsy Seizure : The Journal of the British Epilepsy Association, 17(4), 297-301.

Hjortskov, N., Rissen, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Sogaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work European Journal of Applied Physiology, 92(1-2), 84-89.

Hsu, K. E., Man, F. Y., Gizicki, R. A., Feldman, L. S., & Fried, G. M. (2008). Experienced surgeons can do more than one thing at a time: Effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons Surgical Endoscopy, 22(1), 196-201.

130

Hyltander, A., Liljegren, E., Rhodin, P. H., & Lonroth, H. (2002). The transfer of basic skills learned in a laparoscopic simulator to the operating room Surgical Endoscopy, 16(9), 1324-1328.

I. , S., G., H., M., M., M., T., & J., H. (1997). Effect of respiration and posture on heart rate variability. Physiological Research, 46(2), 173-179.

Ilgen, D. R. (1990). Health issues at work: Opportunities for industrial/organizational psychology. American Psychologist, 45(2), 273 -283.

Innes, L. G., & Allnutt, M. F. (1967). Performance measurement in unusual environments. IAM Technical Memorandum no 298,

J., B., & H., L. (1977). Biofeedback, verbal instructions and the motor skills analogy. , New York: Plenum Press.

James, D., & Richard, S., Lazarus. (1952). The effects of psychological stress upon perceptual-motor performance. Lackland Air Force Base: Perceptual and Motor Skills Research Laboratory, Human Resources Research Center, Air Training Command.

James, E.,Driskell, & Eduardo, S. (1996). In Edwin A.,Fleishman (Ed.), Stress and human performance. New Jersey: Lawrence Erlbaum.

Jay, M., Finkelman, & Carl, K. (1980). An information-processing interpretation of air traffic control stress . Human Factors: The Journal of the Human Factors and Ergonomics Society, 22(5), 567-567.

Jon, R., Pluyter, Sonja, N.,Buzink, Anne-F., R., & Jack, J., Jakimowicz. (2010). Do absorption and realistic distraction influence performance of component task surgical procedure? . Surgical Endoscopy, 24(4), 902-907.

Jowett, N., LeBlanc, V., Xeroulis, G., MacRae, H., & Dubrowski, A. (2007). Surgical skill acquisition with self-directed practice using computer-based video training American Journal of Surgery, 193(2), 237-242.

Kaloupek, D. G., White, H., & Wong, M. (1984). Multiple assessment of coping strategies used by volunteer blood donors: Implications for preparatory training Journal of Behavioral Medicine, 7(1), 35-60.

Kaloupek, D. G., & Stoupakis, T. (1985). Coping with a stressful medical procedure: Further investigation with volunteer blood donors Journal of Behavioral Medicine, 8(2), 131 -148.

Keinan, G., Friedland, N., & Ben-Porath, Y. (1987). Decision making under stress: Scanning of alternatives under physical threat Acta Psychologica, 64(3), 219 -228.

Kivimäki, M., & Lusa, S. (1994). Stress and cognitive performance of fire fighters during smoke-diving Stress Medicine, 10(1), 63 -68.

Kleiger, R. E., Stein, P. K., Bosner, M. S., & Rottman, J. N. (1992). Time domain measurements of heart rate variability Cardiology Clinics, 10(3), 487-498.

Köhn, -.,Seyer, Frauendorf, G., Mahnke, H., H., E. & U., G.,W. (1985). Biologic expenditure in surgically active orthopedists Retrieved 3/29/2012, 2012, from http://scholar.qsensei.com/content/rj830

Korndorffer, J. R.,Jr, Dunne, J. B., Sierra, R., Stefanidis, D., Touchard, C. L., & Scott, D. J. (2005). Simulator training for laparoscopic suturing using performance goals translates to the operating room Journal of the American College of Surgeons, 201(1), 23-29.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1985). Processing of stimulus properties: Evidence for dual-task integrality. Journal of Experimental Psychology: Human Perception and Performance, 11(4), 393 -408.

Krantz, G., Forsman, M., & Lundberg, U. (2004). Consistency in physiological stress responses and electromyographic activity during induced stress exposure in women and men Integrative Physiological & Behavioral Science, 39(2), 105 -118-118.

Krantz, G., Forsman, M., & Lundberg, U. (2004). Consistency in physiological stress responses and electromyographic activity during induced stress exposure in women and men Integrative Physiological & Behavioral Science, 39(2), 105 -118.

Krueger, W. C. F. (1929). The effect of overlearning on retention. Journal of Experimental Psychology, 12(1), 71 -78.

Langelotz, C., Scharfenberg, M., Haase, O., & Schwenk, W. (2008). Stress and heart rate variability in surgeons during a 24-hour shift Archives of Surgery, 143(8), 751 -755.

Larsson, G. (1987). Routinization of mental training in organizations: Effects on performance and well-being. Journal of Applied Psychology, 72(1), 88 -96.

Lazarus, R., S, & Launier, R. (1978). Stress-related transactions between person and environment. Perspectives in Interactional Psychology, , 287-327.

Li, G., Baker, S. P., Grabowski, J. G., & Rebok, G. W. (2001). Factors associated with pilot error in aviation crashes. Aviation, Space, and Environmental Medicine, 72(1), 52-58.

Linn, B. S., & Zeppa, R. (1984). Stress in junior medical students: Relationship to personality and performance. Journal of Medical Education, 59(1),(1), 7-12.

Loewen, L. J., & Suedfeld, P. (1992). Cognitive and arousal effects of masking office noise Environment and Behavior, 24(3), 381 -395.

Logan, G. D. (1985). Executive control of thought and action Acta Psychologica, 60(2-3), 193 - 210.

Lusk, S. L., Gillespie, B., Hagerty, B. M., & Ziemba, R. A. (2004). Acute effects of noise on blood pressure and heart rate Archives of Environmental Health, 59(8), 392-399.

Mathews, K. E., & Canon, L. K. (1975). Environmental noise level as a determinant of helping behavior. Journal of Personality and Social Psychology, 32(4), 571 -577.

McCormack, H. M., de L. Horne, D. J., & Sheather, S. (2009; 1988). Clinical applications of visual analogue scales: A critical review Psychological Medicine, 18(04), 1007.

McDougall, E. M., Corica, F. A., Boker, J. R., Sala, L. G., Stoliar, G., Borin, J. F., et al. (2006). Construct validity testing of a laparoscopic surgical simulator Journal of the American College of Surgeons, 202(5), 779-787.

Meichenbaum, D. H., & Deffenbacher, J. L. (1988). Stress inoculation training The Counseling Psychologist, 16(1), 69 -90.

134

Mercankaya, R. (1974). Estimation of the physiological parameters of heart-rate and oxygen-consumption during heat and work stress.

Miyake, A., & Shah, P. Models of working memory: Mechanisms of active maintenance and executive control Cambridge ; Cambridge University Press, 1999.

Molinas, C. R., Win, G., Ritter, O., Keckstein, J., Miserez, M., & Campo, R. (2008; 2008). Feasibility and construct validity of a novel laparoscopic skills testing and training model Gynecological Surgery, 5(4), 281-290.

Moorthy, K., Munz, Y., Sarker, S. K., & Darzi, A. (2003). Objective assessment of technical skills in surgery BMJ (Clinical Research Ed.), 327(7422), 1032-1037.

Moorthy, K., Munz, Y., Undre, S., & Darzi, A. (2004). Objective evaluation of the effect of noise on the performance of a complex laparoscopic task Surgery, 136(1), 25-30; discussion 31.

Munz, Y., Kumar, B. D., Moorthy, K., Bann, S., & Darzi, A. (2004). Laparoscopic virtual reality and box trainers: Is one superior to the other? Surgical Endoscopy, 18(3), 485-494.

Murata, A. (1992). Effects of respiration on heart-rate variability measures Electronics and Communications in Japan (Part III: Fundamental Electronic Science), 75(2), 107 -116.

Myung, G., Kang, Sang, B., Koh, Bong, S., Cha, Jong, K., Park, Jong, M., Woo, & Sei, J.,Chang. (2004 Oct). Association between job stress on heart rate variability and metabolic syndrome in shipyard male workers Yonsei Med J., 45(5), 838-846.

Neal, E. S., Anthony, G. G., Sanziana, A. R., Michael, K. O., Vipin, K. B., Dana, K. A., et al. (2002). Virtual reality training improves operating room PerformanceResults of a randomized, double-blinded study. Annals of Surgery, 236(4), 458-464.

Newmark, J., Dandolu, V., Milner, R., Grewal, H., Harbison, S., & Hernandez, E. (2007). Correlating virtual reality and box trainer tasks in the assessment of laparoscopic surgical skills American Journal of Obstetrics and Gynecology, 197(5), 546.e1-546.e4.

Novaco, R. W., Stokols, D., & Milanesi, L. (1990). Objective and subjective dimensions of travel impedance as determinants of commuting stress American Journal of Community Psychology, 18(2), 231 -257.

Payne, R. (1986). Heart rate as an indicator of stress in suegeons and anaesthetists Journal of Psychosomatic Research, 30(4), 411 -420.

Prince., C., Bowers., C., A., & Salas., E. '. (1994). Stress and crew performance: Challenges for aeronautical decision making training. In J. N. , M. N. & F. R. (Eds.), Aviation psychology in practice (pp. 286-305). England: Avebury:

R., C.,Aitken. (1969 October). Measurement of feelings using visual analogue scales. Royal Society of Medicine, 62(10), 989-993.

Richard, S.,Lazarus, & Susan, F. (1984). Stress, appraisal, and coping. New York: Springer Pub. Co.

S., S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. ()

Scott, D. J., Bergen, P. C., Rege, R. V., Laycock, R., Tesfay, S. T., Valentine, R. J., et al. (2000). Laparoscopic training on bench models: Better and more cost effective than operating room experience?11No competing interests declared. Journal of the American College of Surgeons, 191(3), 272 -283.

Seyle, H. (1936). A syndrome produced by diverse nocuous agents. Nature, 138, 32.

Strentz, T., & Auerbach, S. M. (1988). Adjustment to the stress of simulated captivity: Effects of emotion-focused versus problem-focused preparation on hostages differing in locus of control. Journal of Personality and Social Psychology, 55(4), 652 -660.

Suess, W. M., Alexander, A. B., Smith, D. D., Sweeney, H. W., & Marion, R. J. (1980). The effects of psychological stress on respiration: A preliminary study of anxiety and hyperventilation Psychophysiology, 17(6), 535 -540.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10(3), 251 -296.

Taelman, J., Vandeput, S., Spaepen, A., & Huffel Chapter 324, S. (2009). 4th European conference of the international federation for medical and biological engineering; influence of mental stress on heart rate and heart rate variability 22, 1366 -1369.

Van Galen, G. P., Van Doorn, R. R., & Schomaker, L. R. (1990). Effects of motor programming on the power spectral density function of finger and wrist movements. Journal of Experimental Psychology: Human Perception and Performance, 16(4), 755 -765.

Van Gemmert, A. W. A., & Van Galen, G. P. (1997). Stress, neuromotor noise, and human performance: A theoretical perspective. Journal of Experimental Psychology: Human Perception and Performance, 23(5), 1299 -1313.

van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006; 2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning Applied Cognitive Psychology, 20(3), 343 -352.

Weitz., J. (1966). Stress. Institute for Defense Analyses, , 251.

Wertkin, R., A. (1985). Stress-inoculation training: Principles and applications. Social Casework, 66(10), 611-616.

Wetzel, C. M., Kneebone, R. L., Woloshynowych, M., Nestel, D., Moorthy, K., Kidd, J., et al. (2006). The effects of stress on surgical performance American Journal of Surgery, 191(1), 5-10.

Wetzel, C. M., Kneebone, R. L., Woloshynowych, M., Nestel, D., Moorthy, K., Kidd, J., et al. (2006). The effects of stress on surgical performance American Journal of Surgery, 191(1), 5-10.

Wetzel, C. M., Kneebone, R. L., Woloshynowych, M., Nestel, D., Moorthy, K., Kidd, J., et al. (2006). The effects of stress on surgical performance American Journal of Surgery, 191(1), 5-10.

Wulf, G., Shea, C. H., & Whitacre, C. A. (1998). Physical-guidance benefits in learning a complex motor skill Journal of Motor Behavior, 30(4), 367-380.

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses Journal of Experimental Social Psychology, 2(2), 160 -168.

Zakay, D., & Wooler, S. (1984). Time pressure, training and decision effectiveness Ergonomics, 27(3), 273-284.

Zapf, D. (1988). Methodological issues in the study of work stress: Objective vs. subjective measurement of work stress and the question of longitudinal studies (pp. 375-411)

ZEIER, H. (1994). Workload and psychophysiological stress reactions in air traffic controllers Ergonomics, 37(3), 525 -539.