4-22-2014 12:00 AM

Validation of a Novel Methodology for Evaluating Surgical Tissue Plane Identification

Syed T. Ali
The University of Western Ontario

Supervisor
C M Schlachta
The University of Western Ontario

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

Recommended Citation

Ali, Syed T., "Validation of a Novel Methodology for Evaluating Surgical Tissue Plane Identification" (2014). *Electronic Thesis and Dissertation Repository*. 2037.
https://ir.lib.uwo.ca/etd/2037

VALIDATION OF A NOVEL METHODOLOGY FOR EVALUATING SURGICAL TISSUE PLANE IDENTIFICATION

(Integrated Article)

by

Syed Ali

Graduate Program in Surgery

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Introduction: The ability to identify tissues planes is a fundamental surgical skill. This study attempts to validate a previously developed test for assessing this ability.

Methods: 48 video captured images from laparoscopic right hemicolectomies were presented on an iPad to 18 surgeons who were grouped based on experience (Consultants (C), Senior trainees (S), and the Junior trainees (J)). Subjects were asked to draw a line indicating the tissue plane of dissection. Lines were compared by a modified Hausdorff measure. Within group variability represented group precision and trainee accuracy was determined from comparison to Consultants.

Results: Within group comparisons demonstrated Consultants to be most precise with statistical significance in 14/25 images. Comparing Seniors and Juniors with Consultants demonstrated Seniors were significantly more accurate than Juniors in 14/22 images.

Conclusion: This tool is sufficiently sensitive to discriminate between surgeons of different levels of experience based on measures of precision and accuracy.

Keywords

Surgical tissue planes, visuospatial, simulation, Hausdorff, colorectal

Co-Authorship Statement

The written material of this thesis is the original work of the primary author Syed Ali with significant contributions from the co-authors.

1) Syed Ali: Syed Ali is responsible for the study concept and design, data acquisition, analysis, interpretation, drafting, final approval and accountability for manuscript. The plan is to submit the integrated article of the thesis to Surgical Endoscopy, which is a peer reviewed surgical journal.

2) Roy Eagleson: Dr. Eagleson was responsible for the study design, analysis and interpretation of data as well as revisions and final approval of the draft.

3) Christopher Schlachta: Dr. Schlachta provided the primary supervisory role to Syed Ali in conducting this study. He was also responsible for the study design and concept, data analysis and interpretation. In addition, he was involved in the final approval and accountability of the manuscript.

Acknowledgements

I would express my sincerest gratitude towards the following people:

1) Christopher Mark Schlachta for being a patient and exemplary supervisor and giving me very helpful advice as well as encouragement at every step of the study.

2) Dr. Roy Eagleson for providing extremely valuable suggestions regarding the mathematical and geometric aspects of our study design.

3) Dr. Tina Mele for her advice and suggestions regarding the final manuscript and also taking time out from her busy schedule to be part of my thesis committee.

4) The entire faculty associated with the Masters of Surgery program, especially Dr. Forbes, Dr. Mele and Dr. Sener for providing me with a sound academic background to be able to successfully complete and conduct this study.

5) Dr. Abdul R. Lawendy for giving me very valuable suggestions and feedback regarding data analysis.

6) Janice Sutherland for all her help as program administrator for our class.

7) My parents and my wife for their patience and support.

Table of Contents

List of Tables

List of Figures

Preface

“There are pleasures that involve no pain or appetite, such as contemplation. Neither practical wisdom nor any state of being is impeded by the pleasure arising from it; it is foreign pleasures that impede, for the pleasures arising from thinking and learning will make us think and learn all the more”.

Aristotle

Chapter 1

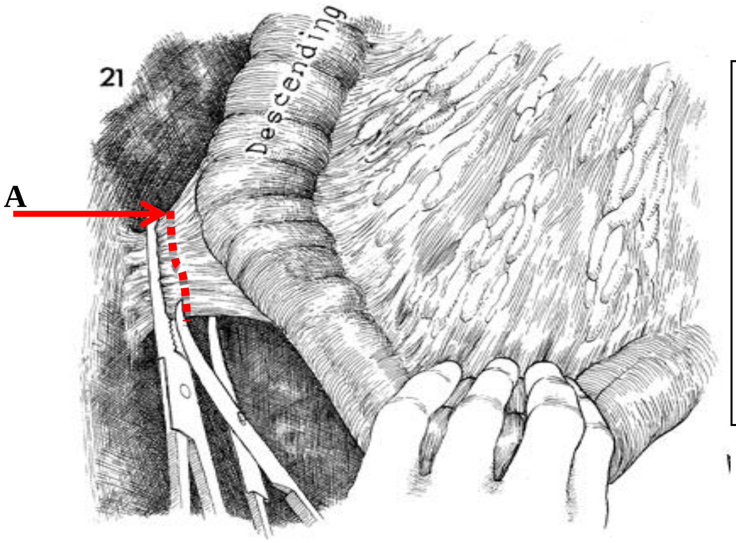
1 Introduction

1.1 Tissue planes of dissection during surgical procedures

Performing surgical procedures requires a great deal of skill across a wide range of levels. It involves a sound knowledge of human anatomy, a comprehensive understanding of the procedure itself, and a certain standard of acceptable technical skill. Mastering surgical skills necessary for a surgical procedure requires expertise in several domains. Anatomical and functional knowledge, capacities for diagnostic reasoning and procedural planning, visuospatial ability (1), and the ability to perform complex sensory-motor tasks in a dynamic and unstructured workspace are all critical to overall performance of the overarching surgical task. Such skills can be refined with repeated practice using the prevalent Halsteadian apprenticeship approach (41, 42) to surgical training.

Most surgical training programs in North America are five years in length. An average surgery resident works more than 85 hours a week (2). Surgical trainees devote a considerable amount of time working directly in the clinical setting in order to gain first-hand experience. This is particularly applicable to their intraoperative exposure. Surgical educators continue to work and research methods to streamline and standardize teaching technical skills to better utilize trainee time and resources. The foremost focus of this research is towards the development of simulation based teaching, where trainees can focus on the development of specific skills required to achieve acceptable competency standards.

One of the most critical portions of teaching a surgical procedure is correct identification of the surgical tissue planes of dissection. Surgical tissue planes of dissection are natural separations between various anatomical structures in the human body. With regards to abdominal surgery, these planes are formed during embryological development of the human body when various internal organs align into their final anatomical position (4). As these developments take place the various organs are separated from each other by zones of fusion forming layers between the various abdominal organs. Identifying the correct tissue plane to separate these layers from each other is of paramount importance for abdominal surgeons, since doing so forms the foundation of separating a particular abdominal organ from the surrounding structures safely in order to remove or maneuver it during an operation. An example of tissue plane of dissection is illustrated in the figure below.



The surgeon has identified the white line of Told's (A) which separates the colon from the reteroperitoneum. Cutting tissue in this correct plane will eventually separate the colon completely from its embryological attachments. (Taken from the Atlas of Pelvic Surgery Online edition) (40)

Figure 1.1: Identification of Surgical Tissue planes

These and the many other planes of dissection encountered by surgeons are not straight uniform lines. Rather they are convoluted and can appear in a complicated series of patterns. Only an expert in this field can claim intuition of these patterns. Staying in the correct tissue plane results in an efficient and safe surgery since these lines of fusion are avascular and do not contain vital structures which need to be preserved during an operation. Critical mistakes and complications occur when the surgeon fails to recognize the correct tissue plane and this increases the likelihood of bleeding and inadvertent tissue injury (5).

1.2 How trainees learn the correct tissue plane

The traditional methods of teaching localization of surgical tissue planes have involved the use of anatomical textbooks and surgical atlases, as well as didactic methods and interactive hands-on teaching in the operating room.

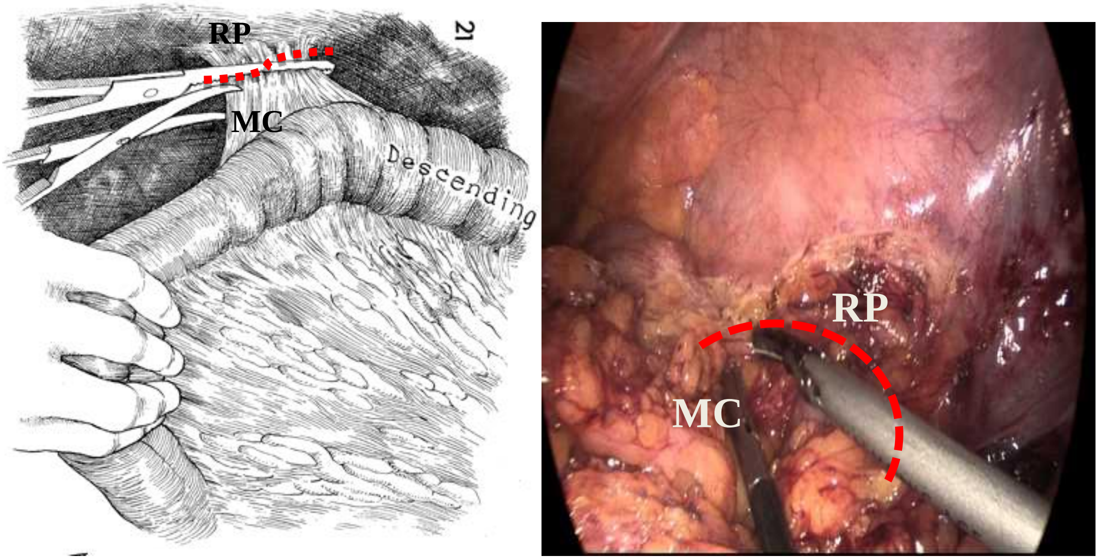


Figure 1.2: Comparing textbook picture to real life intra operative view.
Both images show the same step of the procedure showing the tissue plane between the Reteroperitoneum (RP) and the Mesocolon (MC) highlighted by the red line. Notice the complex visual cues in the real life image on the right (Taken from The Atlas of Pelvic Surgery- Online edition (40)).

Figure 1.2 demonstrates a very simplistic representation of the anatomy highlighting the correct tissue plane in the image on the left. The picture on the right is the same anatomical representation taken from a still captured image from a real laparoscopic case. The reader’s attention is drawn to the extent of visual detail in this picture along with the complex visuospatial arrangement of different structures as compared to its simplistic representation in the textbook image (47, 48).

The practical method of learning to identify correct tissue planes lies in the traditional apprenticeship system of learning operative skills. The trainee sees the anatomy or structures in the operating room, interprets the visual information presented to him or her and then responds by cutting where they estimate the location of the tissue plane to be. They are given feedback immediately if they are not correct by the surgical instructor. With repetition of this process, it is expected that the trainee will ultimately learn how to reason about the visual scene and estimate the location of the tissue plane. This method of education, although more effective than didactic learning, still requires a considerable amount of time, effort and has an unpredictable learning curve (49, 35). Since the literature clearly demonstrates that skills requiring visuospatial ability (VSA) can be enhanced with practice, identification of the correct surgical dissection plane is an area which is ripe for simulation based education (6, 7). The authors of this study propose that simulation can act on the interaction between feedback and response and has the potential to improve educational experience for surgical trainees. Ultimately, the goal of this study is to design a meaningful simulation based educational tool to facilitate surgeons’ tissue plane identification.

1.3 Task Analysis and Context Integration

In order to develop an effective teaching tool to facilitate surgical tissue plane identification, one must first try to understand the nature of skills required for accomplishing the overarching surgical procedure in an effective way to assess and compare the surgeon’s ability to perform the task. Surgical educators and researchers have tried to quantify surgical technical skill such that its assessment and comparison can be standardized and applied to a broad range of surgeons. The challenge in this case is the complexity and broad nature of surgical tasks, and the difficulty of acquisition of surgical skills. Ringsted *et al.* (8) have shown that the specific content of the task including the kind and complexity of the procedure serve as important factors involved in the acquisition of technical skills required by surgeons.

Tissue plane identification during an operation is a specific task which involves interpreting complex visual stimulus within variable anatomical arrangements of human organs. The plane of dissection tends to lie between the various anatomical structures (Figure 1.3).

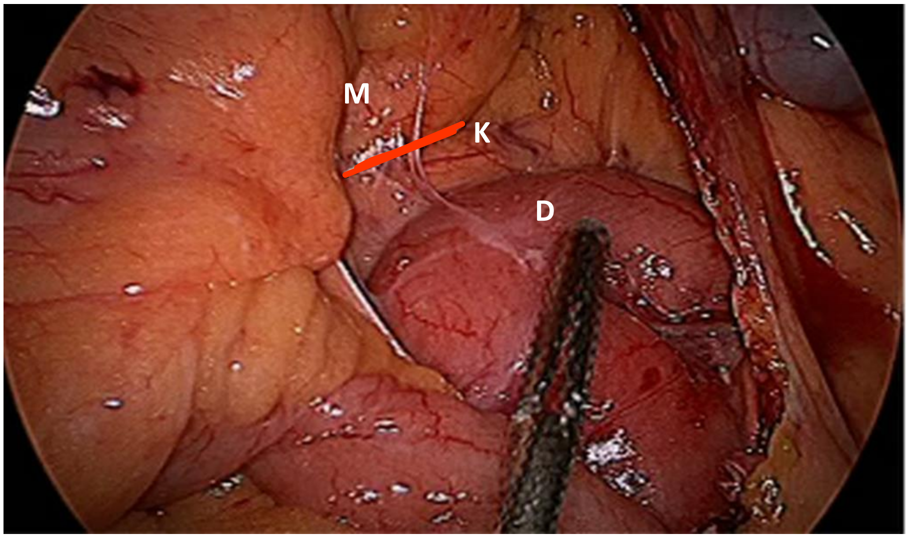


Figure 1.3: Intraoperative view of a right hemicolectomy.
The goal is to separate the Mesocolon (M) from the Kidney (K) by staying and dividing in the correct tissue plane indicated by the red line. The surgeon has to appreciate the anatomical relations, distances and 3 dimensional visuospatial cues.

The surgeon requires the ability to interpret the visual stimulus. Correct interpretation also involves situational awareness towards the three-dimensional relationships between anatomical structures. Thus an individual’s innate visuospatial ability can play a major role in this performance (6). Surgeons' experience tends to play a role as well, since experienced surgeons in actual clinical settings have a better understanding of the location of the tissue plane. Therefore it would seem that experienced surgeons can make use of their visuospatial skills and spatial memory during this task (18).

Although it would seem that visualizing the tissue plane primarily involves the use of visuospatial skills, issues like manual dexterity and tactile sensations might also play a role in this process, since the surgeon needs to know how to manipulate tissues correctly by grasping or applying force on the tissues to expose the tissue plane. In other words the task is a sensorimotor interaction.

The precise extent of the involvement of visuospatial ability and its integration with other human cognitive abilities like manual dexterity, human intelligence and other psychomotor skills is the subject of a number of studies (19). The human intelligence model proposed by Caroll (45) has widely been accepted in human psychology. According to this model human intelligence includes three components: verbal ability, nonverbal reasoning ability, and spatial ability (45). These components are partially separable but are not completely independent. This study chooses to focus on one important task: that of surgical tissue plane identification. The visuospatial component of tissue interaction comes into play when the trainee looks at the surgical field, visuospatially processes what he/she sees, and comes to a conclusion about where the tissue plane is located. Furthermore, the authors of this study propose that trainee performance in this sub-task can be measured and possibly learned separate from motor skills. The focus of this research is to assess and quantify the use of a surgeon's visuospatial ability in identifying tissue planes, which is the first step towards ultimately developing a teaching tool to facilitate this skill.

1.4 What is visuospatial ability and what is its role in tissue plane identification?

To put it in simple terms visuospatial ability means to visually perceive objects and their spatial relationships in the surrounding environment.

It is visuospatial ability which allows humans to retrace their way across the city because of having a visual map in their memory from repeated trips in the past. It allows humans to know that a car is closer to them and smaller than the building just behind the car. Visuospatial skills include a wide variety of individual skills that vary from recognizing brightness/darkness, identifying complex intersecting angles and curves to recognizing faces from the shape of eyes, noses, mouths and hair. It also involves complex human tasks like using maps, solving geometry questions, and recognizing two dimensional representations of three-dimensional objects (9). Visuospatial functions represent the brain's highest level of visual processing, and require the proper functioning of the parietal cortex. This is a complex human skill which is not composed of one construct but in fact is subdivided into various different domains of abilities (9).

Oliveira (10) draws attention to the fact that while spatial ability is a term which is used frequently by psychologists and human factor design experts, there are contradictions in what the different constituents of visuospatial ability are and what roles they play. Many times researchers use the same description under different abilities and vice versa. There is also a general disagreement on the number of components of spatial ability and the exact contents of a component.

One of the most comprehensive works in this field was done by Carroll (11, 45) who did a meta-analysis of more than 140 datasets and detected five major clusters or subdomains which collectively constitute visuospatial ability. Further studies done by Hegarty, Waller and Halpern (12, 13, and 43) added three more to these five domains. What should be emphasized is the lack of general agreement in the actual number and specific definitions of each of these domains. Some of these domains are listed below (20, 43, 45):

- Visual Processing (Gv): The ability to perceive, analyze, synthesize, and think with visual patterns, including the ability to store and recall visual representations.

- Visualization (Vz): The ability to apprehend, encode and manipulate visuospatial representations, often involving rotation in two or three dimensions.

- Spatial relations (SR): The speed of manipulating simple visuospatial representations by transformation.

- Closure speed (CS): The speed of retrieving visuospatial representations from long-term memory when presented with incomplete, disguised or obscured forms of those representations.

- Closure flexibility (CF): The speed of identifying given visuospatial patterns in a complex visual environment.

- Perceptual speed (P): The speed of making correct comparisons when given a number of alternative patterns (11).

- Dynamic Spatial Ability (DSA) or Spatiotemporal Ability (SA): Judgments regarding a moving stimulus (12).

- Environmental Ability (EA): Integrating spatial information about natural and artificial objects and surfaces in an individual’s surroundings. These abilities are considered essential for way-finding and navigation (14 and 15).

Figure 1.4 further explains some of the above definitions.

Figure 1.4: Various domains of visuospatial ability (6) [Figure 1.4 removed because of unavailability of copyright permission, original reference in references section (6)].

1.5 Measurements of Visuospatial Ability (VSA)

Human behavioral experts have developed several different performance based tests to measure an individual's visuospatial ability. Figures 1.5 A, B, C and D illustrate some examples (16, 17).

1) The paper folding test (Figure 1.5 A): The subject must imagine that a sheet of paper has been folded in a certain way, a hole is punched through all thicknesses of the paper at a certain point, and the sheet is unfolded. The folding and punching are indicated on the left side of the vertical line, and the subject must select which of the five unfolded sheets on the right of the vertical line is the result.

2) The mental rotations task (Figure 1.5 B): The subject must imagine rotating three-dimensional block figures. The target/criterion figure is represented on the far left, and the subject must determine as quickly and accurately as possible which two of the four option figures on the right are rotations of the target figure.

3) The hidden figures test (Figure 1.5 C): For each pair of figures, the complex pattern on the right includes the simple geometric pattern drawn on the left. The participant must recognize it and pencil it in (highlighted patterns in each complex pattern).

4) The space relations test (Figure 1.5 D): This test measures the ability to visualize a three-dimensional object from a two-dimensional pattern and to visualize how this object would look if rotated in space. It assesses the ability to "think in three dimensions." Subjects would be asked to identify which figure out of the four options (F, G, H, I) would result from the pattern on the left.

Figure 1.5 A, B, C and D: Different tests to assess visuospatial ability. [Removed because of unavailability of copyright permission, original reference in References section (16, 17)].

Detailed analysis of the different tests of VSA suggest that they have a least two aspects in common – each seems to require the execution of a series of mental transformations in two or three dimensions, and in each, intermediate products must be stored temporarily in visuospatial working memory during the processing of other information (11,13). For example, in the mental rotations task, two or more of the block figures must be rotated in order to determine whether the blocks are rotations of the target. Furthermore, the orientations of various parts of a block have to be remembered while other parts are rotated (11, 13).

1.6 Differences in visuospatial ability and genetic links

Human beings tend to differ in their visuospatial ability as has been demonstrated by their variability in performance on the tests mentioned above and a battery of other similar tests checking various aspects of visuospatial perception (18). Individuals who score higher on these tasks are not only faster, have more working memory resources for storing and processing visuospatial information, but also tend to adopt more efficient strategies for solving visuospatial problems.

Environmental, cultural, and social factors as well as gender have been known to affect the development of visuospatial skills in children but the exact influence of all of these factors tends to be a controversial topic (19).

1.7 Literature foundations of visuospatial ability in surgery

While it is not known which sub constructs of human visuospatial ability affect the skills to operate in the correct tissue plane and up to what extent, it is well known that visuospatial ability plays a very important part in the acquisition and practice of surgical technical skills particularly trainees.

One of the first studies in this subject was done by Shuneman *et. al* (23). One hundred and twenty general surgery residents were tested with a neuropsychological test battery and then rated by attending surgeons on surgical skills exhibited during the course of 1445 surgical procedures. The battery of tests included the Minnesota paper folding test (test similar to the example in figure 1.5 A) and a hidden figures test (Figure 1.5 C). The investigators found a statistically significant correlation between general surgical ability of the trainee and performance on the paper folding test but not on the hidden figures test. They also found a surprisingly negative correlation with usual markers of academic excellence (MCAT). As would be expected, the experience of the trainee was strongly correlated with surgical performance. This begs the question of how important visuospatial ability is as surgical experience increases. The authors did not address if a difference existed among trainees with similar experience and whether that difference could be attributed to differing visuospatial skills. Anastakis *et al.* (24) in their review of this study correctly pointed out the lack of a rationale for choosing the particular battery of tests. Also the fact that residents with better technical skill scored higher on one of the tests (Paper folding test) but not on the other was not explained. Thus it can only be speculated what aspect of VSA correlates with improved performance in the operating room. Also measuring surgical abilities on 1445 different surgical procedures is a very heterogeneous group of tasks which would have required a multitude of abilities and skills not only visuospatial but also others. It is hard to conclude which domains of visuospatial ability were checked and correlated with better surgical performance of a particular set of skills. The surgical skill is rated with a comprehensive but subjective checklist and there is no measurable variable.

Murdoch *et al.* (25) conducted a study which was relatively specific in terms of the participant characteristics, task involved and VSA test used. They tested microsurgery trainees on their ability to perform a microsurgical anastomosis and tested their VSA by a space relations test which involved perception of a three-dimensional representation from a two-dimensional image. The investigators found that trainees who got better ratings also scored higher on the space relations test. This gave some weight to the hypothesis that visuospatial ability specifically pertaining to three-dimensional interpretation of two-dimensional images positively correlates with performance on a specific microsurgical task. The surgical ratings received by trainees when actually performing the surgical task (i.e. the microsurgical anastomosis), however, did not have an objective measure associated with them. Also, what aspect of the task was specifically related to visuospatial ability cannot be ascertained with confidence.

Gibbons *et al.* (26) did a similar study on trainees as Scheuneman *et. al.* and found a positive correlation between average rating of technical skills and performance in the hidden figures test (Figure 1.5 C). Again the ratings of technical ability were subjective and what particular aspect of technical skill positively correlated with the hidden figures test was unknown.

An interesting aspect of VSA and surgical skills is the effect it has on experience of the surgeon. This was demonstrated by a study done by Keehner *et al*. (28). The study was done on junior surgical trainees and experienced surgeons who were attending a simulation course. The experience of the surgeon was gauged by the number of procedures done and the technical skill was assessed based on a previously validated scale while performing procedures on a cadaver (29). The VSA was measured by the Paper folding test (Figure 1.5 A). In the junior surgeons, surgical skill and VSA strongly correlated with each other while no such difference was appreciated in the experienced surgeon group. This according to the authors was consistent with the findings of skill acquisition researchers (30), who have shown that cognitive abilities such as spatial ability are important during the initial phase of learning a new skill, but less important in later phases in which skills become increasingly proceduralized.

Keehner *et al.* did another study on this subject using a different set of skills (27). This time they used the ability to drive a 30 degree laparoscopic camera by novices and also allowed them to practice their skills on this task in a simulation based environment for a total of 12 learning sessions. This skill is technically a complex visuospatial task and involves three-dimensional interpretation of two-dimensional views. The authors tested all participants on general reasoning ability and VSA at the beginning of these sessions. General reasoning ability was assessed by the Differential Aptitude Test battery as described by Bennet *et al.* (32). The test comprised sequences of geometric figures with elements changing systematically according to a logical rule. Mental Rotation Test (Figure 1.5 B) and Guay's Visualization of Views Test (31) were used to check for VSA. In Guay's Visualization of Views Test, participants saw a depiction of a 3D object in the center of a cube. The same object from a different viewpoint was depicted below the cube. The task was to indicate the corner of the cube from which the new view was taken. Both the VSA and general reasoning were correlated with the ability to drive a 30 degree laparoscopic camera at the beginning of the training sessions but at the end of the sessions only VSA still correlated with better performance. In fact general reasoning and VSA had a positive correlation amongst each other and when this correlation was eliminated out of the analysis, the effect of VSA became more pronounced, particularly towards the last of the sessions as the experience of the participant increases. The researchers also showed that with repeated practice the participants with low VSA scores also had comparable performance to expert laparoscopic surgeons which were tested in a separate cohort. This study has an interesting result and contrary to earlier held beliefs, shows that even after gaining acceptable proficiency with repeated practice, visuospatially gifted individuals continue to perform better in complex visuospatial tasks (Figure 1.6). The author's hypothesis on this is that the changing correlations reflect a shift from a reliance on strategic or executive processes as assessed by abstract reasoning to exclusively spatial transformation processes. Thus with practice, the strategic component of performance on the laparoscope task decreases, whereas its dependence on the ability to maintain and transform spatial information persists. Besides the fact that this was a purely simulation based environment, one of the other drawbacks of this study is the paucity of long term assessment of these participants, since it is not known if the effect would have continued to persist between VSA and performance on the 30 degree camera if practice even after the 12th training session would have continued.

Figure 1.6: Performance with a laparoscope as impacted by general and spatial ability [figure removed because of copyright, original source in reference (27)].

Another study which highlights the importance of practice in enhancing performance on VSA tasks and differentiating how visuospatial ability affects performance was done by Wanzel *et al.* (33). The authors compared performance on Z-plasty, a spatially complex surgical procedure, to performance on six standardized visuospatial tests. These tests ranged in complexity from relatively simple two-dimensional items such as the snowy pictures test (16) to more complex three-dimensional visualization items such as the form board (16) and mental rotation tests (Figure 1.5 B). They found that only the latter two tests predicted performance on the surgical procedure, which they interpreted as evidence for the involvement of three-dimensional visualization processes in the surgical task. The correlations were strongest for the most spatially complex surgical procedures. Furthermore, only participants who scored higher on visuospatial tests were able to successfully transfer their learning to a more complex version of the Z-plasty procedure. One possible explanation of these findings is that processes such as visualization, mental rotation, and spatial orientation help to support and maintain a mental model of anatomical structure during surgical procedures, and help formulate an end product in mind before the procedure is started. The authors also showed that with practice, subjects who scored low in the VSA tests ultimately were able to perform the Z-pasty at an acceptable standard.

It seems that studies focussing on aspects of VSA pertaining to three-dimensional interpretation reveal the strongest correlation between task performance and innate VSA. Vlez *et al*. (34) did a study on healthy volunteers and found that individuals who score highly on VSA tests also perform better at a computer based visualization battery of tests which involves interpreting three-dimensional information from two-dimensional representation of objects. They also showed that subjects with higher spatial ability had less difficulty with object complexity and hidden properties of an object.

These effects would perhaps play a very important role in surgery and specifically laparoscopic abdominal surgery, which involves working in a two-dimensional environment in a limited space and visually interpreting three-dimensional information. This would also mean that individuals with high visuospatial ability would perhaps perform better in such environments. Also in surgery, there is a degree of variability in performing the same procedure on a different patient—an environment similar to air traffic control where it has been shown that the phenomena of automation of learning does not improve performance to such an extent so as to nullify the effect of visuospatial skills (35).

While most studies point towards a positive correlation between VSA and surgical performance, some studies show no such correlation. Deary *et al.* (36) rated 22 surgical trainees on operating ability and assessed them with a battery of tests on visuospatial ability, intelligence, and personality. There were no significant correlations between surgical ability and visuospatial ability. Thus, these findings contradict earlier studies like Sheuneman *et.al.* (23). Again, this was a heterogeneous sample, with very general ratings of surgical ability as the dependent variable, and little rationale provided for the choice of the specific spatial ability tests used. Thus, the same limitations exist in interpreting both this study and the study by Scheuneman *et.al* (23). It also has been shown that expert surgeons as a group are not exceptional visuospatially (8, 10). Sidhu *et al.* (37) did a study on expert vascular surgeons and novices in assessing their ability of interpretation of three-dimensional structures from two-dimensional endovascular images. They found that perception of overall surface contours of three-dimensional structures from two-dimensional angiographic images was affected by experience and training and was not related to innate visuospatial ability. Contrary to studies by Keehner, Wanzel and Vlez (27, 28, 33, 34), this study does not support the hypothesis that innate VSA affects three-dimensional interpretation of objects.

A rather comprehensive systematic review was done by Maan *et al*. (50). This review explored predictors of better surgical skill by analyzing a total of 27 studies out of which 13 showed VSA to be a predictor of better surgical performance.

1.8 Summary of literature review

Although the above-mentioned studies show some conflicting findings, and many lack specificity and objectivity in what is being tested, certain conclusions can be fairly drawn.

1) VSA does tend to play a role in the acquisition and practice of technical skills in the operating room. These skills play an important role in novices. There is some evidence (27, 33) that they continue to play some role even with increasing experience.

2) Subjects with low VSA are also able to perform at acceptable level of competence with repetitive practice but take longer and more practice to get to that level.

3) The clinical relevance of point 2 is debatable. For educators an important question is whether this variability in time to train and final performance has practical significance in the context of real surgery.

The fact that practice causes improvement in ability introduces the topic of simulation in surgical education which is the focus of this thesis as will be elaborated in the next few chapters.

While there is no methodology available to the authors' knowledge which has focused on developing a simulation based teaching tool to facilitate tissue plane identification by trainees, there is considerable literature that suggests that simulation based surgical education does enhance the educational experience of trainees as pointed out earlier (27,33). Dawe *et al.* (44) reviewed a total of 12 randomized controlled trials involving surgical simulation and operating room performance and found statistically significant improvement in trainee performance as a result of simulation on a number of procedures. An extensive meta-analysis was done in this regard by Sutherland *et al*. (33) who reviewed 30 randomized controlled trials with more than 700 participants. This meta-analysis found mixed role for simulation in surgical education, but noted the small volumes and under powering of a number of trials along with the presence of multiple confounding variables. Most of these studies lacked specificity in explaining which subsets of surgical skills were checked and enhanced while participating in simulation. Still most randomized controlled trials in this review showed that both simulator and standard training groups improved significantly from baseline. Participants' final scores usually did not show differences between the simulator and standard training. The authors concluded from this review that simulation based training might be as good as standard training and has the potential of substituting it in this age of increasing patient safety and accountability. Anastakis *et al.* (24) suggested that in order to understand visual perception and its role in surgical education, subjects need to be assessed on specific tasks and with specific tools. The authors of this study believe that in studying a trainees understanding of the surgical tissue planes they are attempting to analyze a specific skill set of surgery.

1.9 Challenges

There are significant numbers of challenges when attempting to assess this ability and developing a teaching tool. While some surgical simulation devices do assess a participant’s ability to stay in correct tissue planes as part of a simulation's procedure based assessment (PBA) (38), the criterion used to assess these abilities are often vague and ill-defined. Thus, this research tends to be of exploratory nature and there appears to be the lack of a pre-set criteria or methodology to follow in order to assess a trainee’s ability to identify correct tissue planes.

The other challenge is defining a gold standard with which to compare. Surgical tissue planes are approached differently by experts. Thus, one consultant surgeon would approach the surgical plane from medial to lateral whereas the other from the opposite direction. Even more importantly there is often disagreement on the exact location of the tissue plane itself among expert surgeons. There is also a lack of comprehensive task analysis of surgical tissue planes. From the literature review it can be concluded that VSA tends to play a major role in this process and a surgeon’s experience does enhance their use of such skills; but what other cognitive and procedural skills are involved and to what extent is speculation. One major factor might be the use of other psychometric abilities to improve a surgeon's visualization of tissue planes, including pushing and pulling the tissue in the right direction. Hence seeing the tissue plane is a complex ability which might involve several domains of human intelligence and several subsets of VSA. Therefore the first step towards facilitating this learning process is to develop an assessment tool which gauges a surgeon's ability to identify tissue planes. It is only when the ability of a surgeon to identify tissue planes can be quantified then one can develop meaningful simulation. Such an assessment tool should be able to discriminate a surgeon based on their ability to identify a tissue plane. This is based on a prior hypothesis that the most experienced surgeons (consultants) are better than the trainees in identifying the tissue planes, since it is known that they do these operations independently, safely and with an acceptable complication rate. Therefore, the challenge boils down to develop a quantitative assessment of this ability of identifying tissue planes.

1.10 Assessment tool

The authors have devised a novel methodology to derive meaningful information from surgeon-drawn lines highlighting the tissue planes on still captured images from real surgeries. This methodology was used to conduct a pilot study (39). This study showed statistically significant differences in accuracy and precision of participants' ability to identify surgical tissue planes based on surgeon's experience. Thus, more experienced surgeons performed better on this teaching tool.

1.11 Hypothesis

Applying this novel methodology for assessing surgeons' ability to identify tissue planes during laparoscopic right hemicolectomies will demonstrate a correlation between surgeons' experience level and both accuracy and precision, providing content validity to this tool.

1.12 Pilot study

A total of 16 still images were captured from a single laparoscopic rectal cancer case. The images were selected by a surgeon with more than 10 years of experience with this procedure. These images were shown to 12 participant surgeons of variable experience and the participants were asked to draw a line where they believed the tissue plane of dissection existed. The participants were divided into junior trainees, senior trainees and consultants based on their level of experience. Once the lines were drawn, a distance based metric (discussed later) was used to give each line a numerical value.

The lines within a group were compared amongst each other to get intragroup precision values. The results showed that the consultants were the most precise group and these differences reached statistical significance in nine out of the 16 images on a one-way ANOVA. Since there was no gold standard, the consultants were used as the gold standard and the junior and senior trainee lines were compared to the consultant lines as a measure of accuracy. The seniors were more accurate than the juniors as a group ($P<0.05$ in 10/16 images). It was concluded that this methodology represented the foundations of an assessment tool which may reliably distinguish surgeons based on their ability to identify tissue planes. This assessment tool was able to distinguish a surgeon with more experience on measures of precision and accuracy.

1.13 Aims and Objectives

The key objectives of this study are:

1. To generate a simulation tool consisting of a library of images from laparoscopic colon surgery suitable for testing trainees.

2. To validate this simulation tool for assessment on a panel of surgeons of different levels of experience.

3. To develop a library of images with the most discriminative value to be used in the future for further studies and ultimately as a teaching tool.

1.14 References

1. Gallagher AG, Leonard G, Traynor OJ. Role and feasibility of psychomotor and dexterity testing in selection for surgical training. ANZ J Surg. 2009 Mar; 79(3):108-13.

2. American Academy of Graduate Medical Education: Recommendation of Committee on work hour restrictions (Internet Article). AAGME, Chicago, IL (Date cited 02/04/2012). Available from https://www.acgme.org/acgmeweb/tabid/271/GraduateMedicalEducation/DutyHours.asp

3. Schueneman AL, Pickleman J, Hesslein R, Freeark RJ. Neuropsychologic predictors of operative skill among general surgery residents. Surgery. 1984 Aug; 96(2):288-95.

4. Sadler, T W (Thomas W.), Langman, Jan. Langman's Medical Embryology 11th ed. Philadelphia:Wolters Kluwer Lippincott Williams & Wilkins; 2010.

5. West NP, Morris EJA, Rotina O, Cairns A, Finan PJ and Quirke P. Pathology grading of colon cancer surgical resection and its association with survival: a retrospective observational study. Lancet Oncol. 2008; 9: 857–65.

6. M. Keehner, Y. Lippa, D.R. Montello, F Tendick, M. Hegarty. Learning a Spatial Skill for Surgery: How the Contributions of Abilities Change With Practice. Appl. Cognit. Psychol. 2006; 20: 487–503. DOI: 10.1002/acp.1198

7. Kyle R Wanzel, Stanley J Hamstra, Dimitri J Anastakis, Edward D Matsumoto, Michael D Cusimano. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet. 2002; 359: 230–231.

8. Ringsted C, Skaarup AM, Henriksen AH et al. Person-task context: a model for designing curriculum and in-training assessment in postgraduate education. Med Teach. 2006; 28:70–76.

9. Linn, M., & Petersen, A.C. Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development. 1985; 56: 1479-1498.

10. Oliviera T. Dynamic spatial ability: an exploratory analysis and a confirmatory study. International Journal of Aviation Psychology. 2004; 14(1): 19-38.

11. Carroll JB. Abilities in the domain of visual perception. In: Carroll JB, editor. Human Cognitive Abilities: A Survey of Factor-Analytic Studies. 1st Ed. NewYork, NY: Cambridge University Press; 1993. P 304-310.

12. Halpern D. Sex differences in cognitive abilities (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates; 2000. 420 p.

13. Hegarty M, Waller D. Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds). The Cambridge handbook of visuospatial thinking. New York, NY: Cambridge University Press; 2005. P 121-169

14. Allen, G. L Functional families of spatial abilities: poor relations and rich prospects. International Journal of Testing. 2003; 3(3): 251-262.

15. Bell, S., & Saucer, D. Relationship among environmental pointing accuracy, mental rotation, sex, and hormones. Environment and Behavior 2004; 36(2): 251-265

16. Elkstrom *et. al.* Manual for kit of factor referenced cognitive tests. Princeton, NJ : ETS; 1978

17. Vanderburg and Kuse *et al.* Mental rotation: A group test of 3 dimensional spatial visualisation. Perceptual and motor skills. 1978; 69: 915-921

18. Salthouse, T.A., Babcock, R.L., Mitchelle, D.R.D. & Skovronek, E. Sources of individual differences in spatial visualization ability. Intelligence. 1990; 14:187-230.

19. Carolyn B. Mervis, Byron F. Robinson, and John R. Pani. Visuospatial Construction. Am J Hum Genet. Nov 1999; 65(5): 1222–1229 Published online Oct 6, 1999.PMCID: PMC1288273

20. Nguyen, Ngan T., "Anatomy: The Relationship between Internal and External Visualizations" (2012). University of Western Ontario - Electronic Thesis and Dissertation Repository. Paper 468.
http://ir.lib.uwo.ca/etd/468.

21. Rose RJ, Harris EL, Christian JC, Nance WE. Genetic variance in nonverbal intelligence: data from the kinships of identical twins. Science. 1979; 205:1153–1155

22. Tambs K, Sundet JM, Magnus P. Heritability of WAIS subtests: a study of twins. Intelligence. 1984; 8:283–293.

23. Scheuneman AL, Pickleman J, Hesslein R, Freeark RJ. Neuropsychologic predictors of operative skill among general surgery residents. Surgery. 1984; 96:288 –295.

24. Dimitri J. Anastakis, Stanley J. Hamstra, Edward D. Matsumoto. Visual-Spatial Abilities in Surgical Training. Am J Surg. 2000;179:469–471

25. Murdoch JR, Bainbridge LC, Fisher SG, Webster MH. Can a simple test of visual-motor skill predict the performance of microsurgeons? J R Coll Surg Edinb. 1994; 39:150 –152.

26. Gibbons RD, Baker RJ, Skinner DB. Field articulation testing: a predictor of technical skills in surgical residents. J Surg Res. 1986; 41:53–57.

27. Keehner, M., Lippa, Y., Montello, D. R., Tendick, F., & Hegarty, M. Learning a spatial skill for surgery: How the contributions of abilities change with practice. Applied cognitive psychology. 2006; 20: 487–503.

28. Keehner, M. M., Tendick, F., Meng, M. V., Anwar, H. P., Hegarty, M., Stoller, M. L., & Duh,Q. Y. Spatial ability, experience, and skill in laparoscopic surgery. Am J Surg. 2004; Jul 188(1):71-5.

29. Martin J, Regehr G, Reznick RK, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997; 84:273–278.

30. Ackerman PL. Determinants of individual differences during skill acquisition: cognitive abilities and information processing. J Exp Psychol Gen. 1988; 117:288-318

31. Eliot, J., & Macfarlane-Smith, I. An international directory of spatial tests. Windsor, Berks: Nfer-Nelson; 1983. 458 p.

32. Bennet, G. K., Seashore, H. G., & Wesman, A. G. Differential aptitude tests. Sidcup, UK: The Psychological Corporation; 1981.

33. Kyle R Wanzel, Stanley J Hamstra, Dimitri J Anastakis, Edward D Matsumoto, Michael D Cusimano. Effect of visual-spatial ability on learning of spatially-complex surgical skills Lancet 2002; 359: 230–231

34. M. Velez, D. Silver, and M. Tremaine. Understanding visualization through spatial ability differences. IEEE Visualization. 2005; pages 511-518.

35. Jan-Maarten Luursema, Sonja N. Buzink, and J.Jakimowicz. Visuo-spatial ability in colonoscopy simulator training. Adv Health Sci Educ Theory Pract. Dec 2010; 15(5): 685–694. Published online May 9, 2010. doi: 10.1007/s10459-010-9230-y

36. Deary IJ, Graham KS, Maran AGD. Relationships between surgical ability ratings and spatial abilities and personality. J R Coll. Surg. Edinb. 1992; 37:74 –79.

37. Sidhu, R. S., Tompa, D., Jang, R., Grober, E. D., Johnston, K. W., Reznick, R. K., & Hamstra S. J. Interpretation of three-dimensional structure from two-

dimensional endovascular images: implications for educators in vascular surgery. Journal of Vascular Surgery. 2004; 39: 1305-1311

38. Anthony G. Gallagher, Gerald C. O'Sullivan. Fundamentals of Surgical Simulation: Principles and Practice. London, UK: Springer; 2012. p 374.

39. Christopher M. Schlachta, Syed Ali, Hammood Ahmed , Roy Eagleson. A Novel Method for Assessing Visual Perception of Surgical Planes, Manuscript submitted to Surgical Endoscopy.

40. C. Wheeless and M Roenneburg. Atlas of Pelvic Surgery Online, 4th Edition (Internet), Developed by C. Wheeless and M Roenneburg. Maryland. (Date cited 02/04/2014) Available from http://www.atlasofpelvicsurgery.com/10MalignantDisease/17PosteriorExenteration/cha10sec17.html

41. Dobson J, Walker RM. Barbers and Barber-Surgeons of London: A History of the Barbers' and Barber-Surgeons Companies. Oxford: Blackwell Scientific Publications for the Worshipful Company of Barbers; 1979:171.

42. Nguyen L, Brunicardi FC, Dibardino DJ, Scott BG, Awad SS, Bush RL, Brandt ML. Education of the modern surgical resident: Novel approaches to learning in the era of the 80-hour workweek. World J Surg. 2006; 30(6):1120-1127.

43. Tendick F, Hegarty M. Elucidating, Assessing, and Training Spatial Skills in Minimally Invasive Surgery Using Virtual Environments. Proceedings of the AAAI Spring Symposium. 2000; pp.148–155.

44. Dawe SR, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, Maddern GJ. A systematic review of surgical skills transfer after simulation-based training:

laparoscopic cholecystectomy and endoscopy. Ann Surg. 2014 Feb; 259(2):236-48. doi: 10.1097/SLA.245.

45. Schneider, W. J., & McGrew, K. The Cattell-Horn-Carroll model of intelligence. In, D. Flanagan & P. Harrison (Eds.) Contemporary Intellectual Assessment: Theories, Tests, and Issues (3rd ed.) New York: Guilford; 2012. p. 99-144

46. Nguyen,N, Eagleson, R, and de Ribaupierre (2013) "Establishing the validity and reliability of computer-based simulation for cerebral angiography using the ANGIO Mentor Express" The FASEB Journal, Vol 27. Article 959.12

47. van Merriënboer, J. J. and J. Sweller. Cognitive load theory in health professional education: design principles and strategies. Med Educ. 2010; 44(1): 85-93.

48. Grober, E. D., S. J. Hamstra, et al. The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. Ann Surg. 2004; 240(2): 374-381.

49. A N Hopper, M H Jamison, and W G Lewis. Learning curves in surgical practice. Postgrad Med J. Dec 2007; 83(986): 777–779.

50. Maan ZN, Maan IN, Darzi AW, Aggarwal R. Systematic review of predictors of surgical performance. Br J Surg. 2012 Dec; 99(12):1610-21.

2 Methodology

2.1 Overview of study design

This is a prospective study. A total of 48 images taken from various laparoscopic right hemicolectomies done on real patients were shown to a total of 18 participants with different levels of experience. Participants were asked to draw a line to indicate where they thought the tissue planes were located. Lines were analyzed based on the assessment tool developed and utilized during the pilot study (2). Performance of each group was compared for precision. The senior and junior trainee lines were also compared with the consultants to get a measure of accuracy.

2.2 Choice of procedure

The first step towards choosing images for this study was to choose an operation. The pilot study had collected images from a single case of laparoscopic low anterior resection. This is a fairly advanced procedure which requires considerable surgical skills and experience. For the purposes of that study, it held the advantage of providing many opportunities for tissue plane identification. On the other hand, because of the advanced nature of the procedure many trainee participants had minimal experience with this procedure and thus the content validity of the methodology could be questioned i.e. performance on those images will not only differ with the surgeons' experience resulting from the lack of knowledge of the correct tissue plane but also a general sense of unfamiliarity with the anatomical orientation of the procedure.

Based on these considerations, laparoscopic right hemicolectomy was chosen as the procedure to use for this study. This procedure confers the following advantages:

- Broad applicability since the procedure is performed very commonly amongst general surgeons. Trainees get exposed to it at an earlier point of their learning cycle.
- There are several opportunities during the procedure which involve identification of the tissue planes of dissection.
- The procedure is performed by the authors of this study in considerable numbers, thus the authors have access to a large archive of video library for this procedure. This resulted in a large sample of images for testing.

2.3 Determination of sample size

Since this is an exploratory study with no previously done standard to determine an accurate estimate of participant sample size, the pilot study was used to determine sample size as below:

In the pilot study, the mean pairwise distance between lines for consultants (c) was found to range from 366 to 1685 pixels. For trainees (t), the range was 343 to 2917 pixels.

Assuming: $\mu c = 750$ pixels, $\mu t = 1500$ pixels, $\sigma = 500$ pixels with

$\alpha = 0.017$ (Sidak correction for multiple comparisons of three groups)

$\beta = 0.8$

Thus, sample size would be 10 comparisons per group, based on a single image analysis. To obtain 10 pairwise comparisons in a group, the group needs to have at least five participants. One more participant per group was recruited than what was required based on this analysis.

2.4 Ethics approval

An ethics approval was obtained before the start of the study from the ethics board associated with the Western University.

2.5 Selection of images

This study involved capturing a large number of images from a video archive of a total of 12 cases of laparoscopic right colectomy performed by a single surgeon during the times of 2006-2010 at London Health Sciences Center. These videos had been recorded from a high definition laparoscopic camera. These videos were then reviewed on a standard video playing computer software (VLC player, École Centrale Paris) and relevant still images were captured for this study. Relevant images mean images which highlight the steps of the procedure where dissection through the tissue planes is necessary.

This entire process generated a total of 1126 images. An expert panel consisting of the investigators of this study which include Christopher Schlachta, Syed Ali, and Roy Eagleson, then reviewed these images. These individuals represent expertise in advanced laparoscopic surgery training, and specific cognitive and perceptual-motor skills respectively. The panel shortlisted a total of 48 images based on the following criteria:

1) The image represents an operative moment where a tissue plane of dissection should be identified and
2) The ability to identify the dissection plane in these images would likely be of variable difficulty for novices and experts.
3) In selecting an image all three members of the panel had to agree on point 1 and point 2.

There was no prior knowledge of how many images would be required for an overall study score nor how many images will be of discriminating value. The pilot study included 16 images only. The aim in selecting images for this study was to incorporate all the different steps of the procedure which highlight the importance of identifying tissue planes. It was desired that the selected images should represent all the different aspects of the operation in a broad and comprehensive manner but at the same time achieve feasibility of time required by the study subjects to perform on the assessment tool. Because of these reasons the expert panel decided on 48 images.

2.6 Labelling and identification of images

Each selected image was given a unique name which was based on the procedure type, the confidential unique identifier attributed to the patient, the step of the procedure and the timestamp of the captured image from its original video. Thus an individual image's name would be:

RHC72-S6-T01223210

Where RHC= Right hemicolectomy, 72 is the unique video identifier, s6 =step number six of the procedure, T01223210= time stamp from the video capturing the hour, minutes, seconds and frame number from the video. Thus in the above example 1 hour 22 minutes 32 seconds and 10th frame of the video is the exact time location of the still captured image. Information on the unique identifier was kept in a master list and was only accessible to the principle investigator because of patient confidentiality as approved by the REB. Although there is no standard classification which divides a right hemicolectomy, for this research each image was divided and labeled to one of the seven steps:

Step 1: Isolation of the illeocolic pedicle

Step 2: Medial mobilization

Step 3: Separation of small bowel mesentery from reteroperitoneum

Step 4: Lateral mobilization

Step 5: Hepatic flexure mobilization

Step 6: Opening of the lesser sac

Most surgeons agree that the above steps are required while performing a right hemicolectomy via a medial to lateral approach (6).

2.7 Identification of Subjects

Volunteer participants were recruited from surgical colleagues and trainees in the Division of General Surgery at Western University and the University of Toronto. After obtaining informed consent for the study, participants were divided into three groups based on their levels of experience (6 surgeons per group).

1. The consultant group (C) included general surgeons with Royal College certification in general surgery, already in practice at one of the hospitals affiliated with Western University or the University of Toronto. They perform colorectal surgery as part of their clinical practice. This is considered the expert group with which trainees were referenced. To be included in the consultant group a consultant had to have performed more than 50 laparoscopic right hemicolectomies independently. This number is expected to be associated with reasonable technical competence in performing this procedure (8,9,10,11).

2. The senior trainees group (S) were residents in their third to fifth years of general surgery residency at Western University.

3. The junior trainee (J) group were in their first two years of general surgery residency at Western University, which is a Royal College accredited Canadian general surgery residency program.

2.8 Plane identification exercise

Images were transferred to an iPad 2 (Apple Computers) and presented in Sketchbook Pro software (Autodesk Inc. San Rafael, California). Subjects were able to draw on the presented images using a stylus for capacitive touch screens (Slim stylus, Targus, Anaheim, California). Each subject's line was saved in a separate layer for later analysis. For each of the images, the subjects were asked to draw a line clearly and precisely where they saw the tissue plane of dissection.

The images were not labeled but there was a standard set of instructions for each image. The purpose of this was to make sure participants were oriented to the step of the procedure involved and oriented anatomically. Uniform sets of instructions were given to each participant.

2.9 Transfer of images from a tablet to a personal computer for analysis

After data collection, the images were transferred onto a secure personal computer for analysis. Size and pixel quality of every image was kept the same. This was necessary to ensure the accuracy of the analysis.

2.10 Analysis of performance

Once data collection was complete, lines were analyzed by a novel methodology (Figure 2.1). This methodology involved using a distance metric similar to the Hausdorff measure used in computational geometry (3), which is based on the Euclidean distance between evenly-spaced points along the arc.

2.11 Hausdorff distance

Hausdorff distance was proposed by Felix Hausdorff (1868-1942) to measure distances between different sets of objects. Modified versions of this distance metric have broad applications in the fields of geometrics and computer graphics but have also been used in the health industry (5).

This distance metric measures how far two subsets of a metric space are from each other. Instead of just taking the shortest distance between two geometric figures, lines or arcs, it also takes into account the maximum distance possible between two objects. The distance metric used in this study is different from the Hausdorff distance in that it preserves the sequence of pairs of each line in performing a pairwise Euclidean distance calculation. Each line is sampled into a discrete set of points (same number of points for each line) and the distance metric is established by summing the pairwise distances.

Using MS Paint (Microsoft Office Professional Edition, 2011, Microsoft Corporation, Seattle, WA), each line was iteratively bisected three times resulting in eight equidistant spaces which are represented by nine equidistant points (Figure 2.1). Cartesian coordinates (X,Y) are plotted on each of these nine points, thus giving a numerical representation of the lines' location. For each study group (G), this distance metric between any two lines (a, b) for a given image (*i*) was calculated as the summed distance (*d*), in pixels, between each of the (p=9) coordinate pairs (j) as follows:

$$d(Gab\colon i) = \sum_{j=1}^{p} \sqrt{(Gax_j - Gbx_j)^2 + (Gay_j - Gby_j)^2}$$

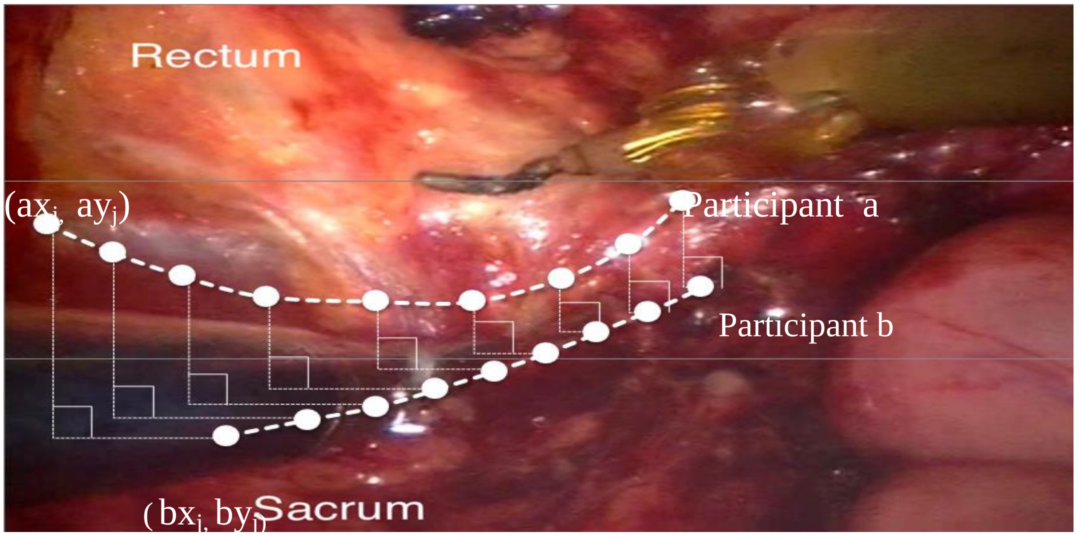


Figure 2.1: Comparing 2 lines with each other.

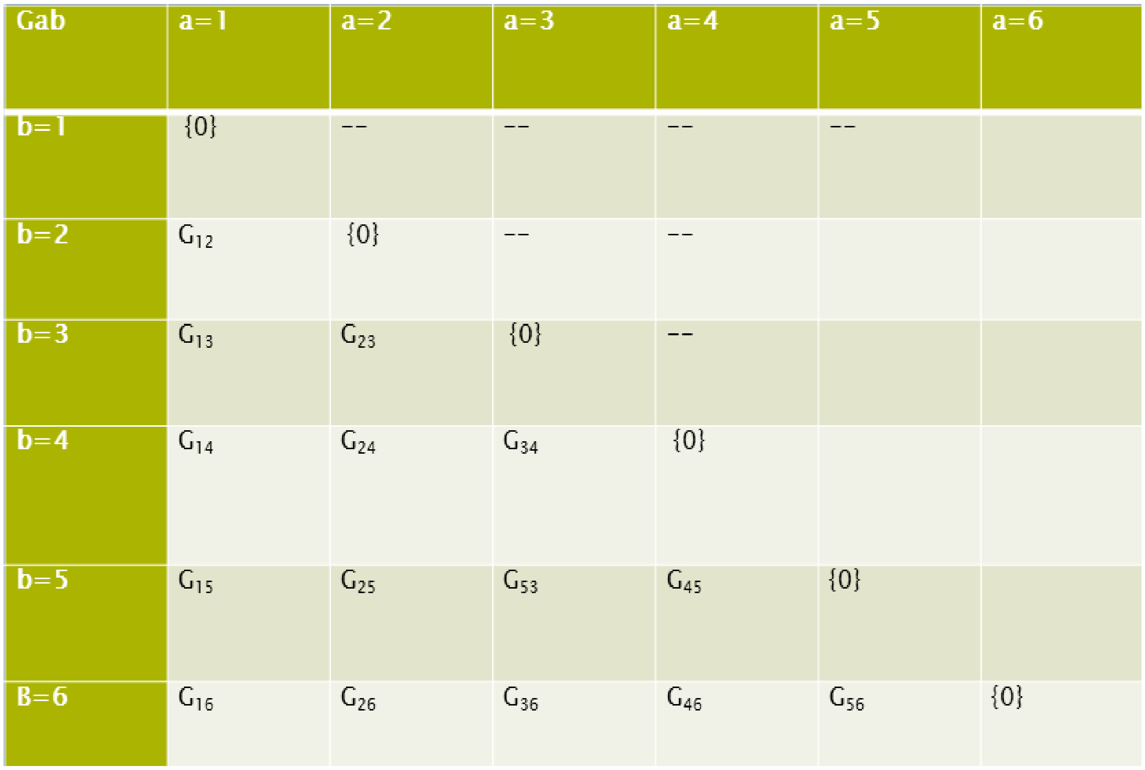
2.12 Order of the Cartesian co-ordinates

While calculating this distance metric, lines were marked with the points in ascending order of their Euclidian distance i.e. Point 1 on a line would be the lowest value of Cartesian coordinates whereas Point 9 would be the highest.

2.13 Analysis of precision

The lines within a group were compared amongst each other for each image to get intragroup precision values. The smaller the values of the distance metric, the closer the lines are to each other within that group, and the more precise the group is. Since there are six participants within a group and each member of a group is being compared with every other member of the same group, there are 15 comparisons per group per image (Figure 2.2).

Mean of these 15 values was calculated to get an estimate of the precision of the particular group for each image. For each image, precision values of the three groups were compared using one way analysis of variance (ANOVA). Where statistically significant differences were found, pairwise comparisons were performed with t-tests (2-tailed unequal variance) with level of significance corrected for multiple comparisons (Sidak’s correction: $p<0.017$). These steps are summarized in figure 2.3.



Gab	a=1	a=2	a=3	a=4	a=5	a=6
b=1	{0}	--	--	--	--	
b=2	G_{12}	{0}	--	--		
b=3	G_{13}	G_{23}	{0}	--		
b=4	G_{14}	G_{24}	G_{34}	{0}		
b=5	G_{15}	G_{25}	G_{53}	G_{45}	{0}	
B=6	G_{16}	G_{26}	G_{36}	G_{46}	G_{56}	{0}

Figure 2.2: Obtaining the number of precision comparisons per group.

When six participants are compared to each other, 15 comparisons of precision per group are obtained.

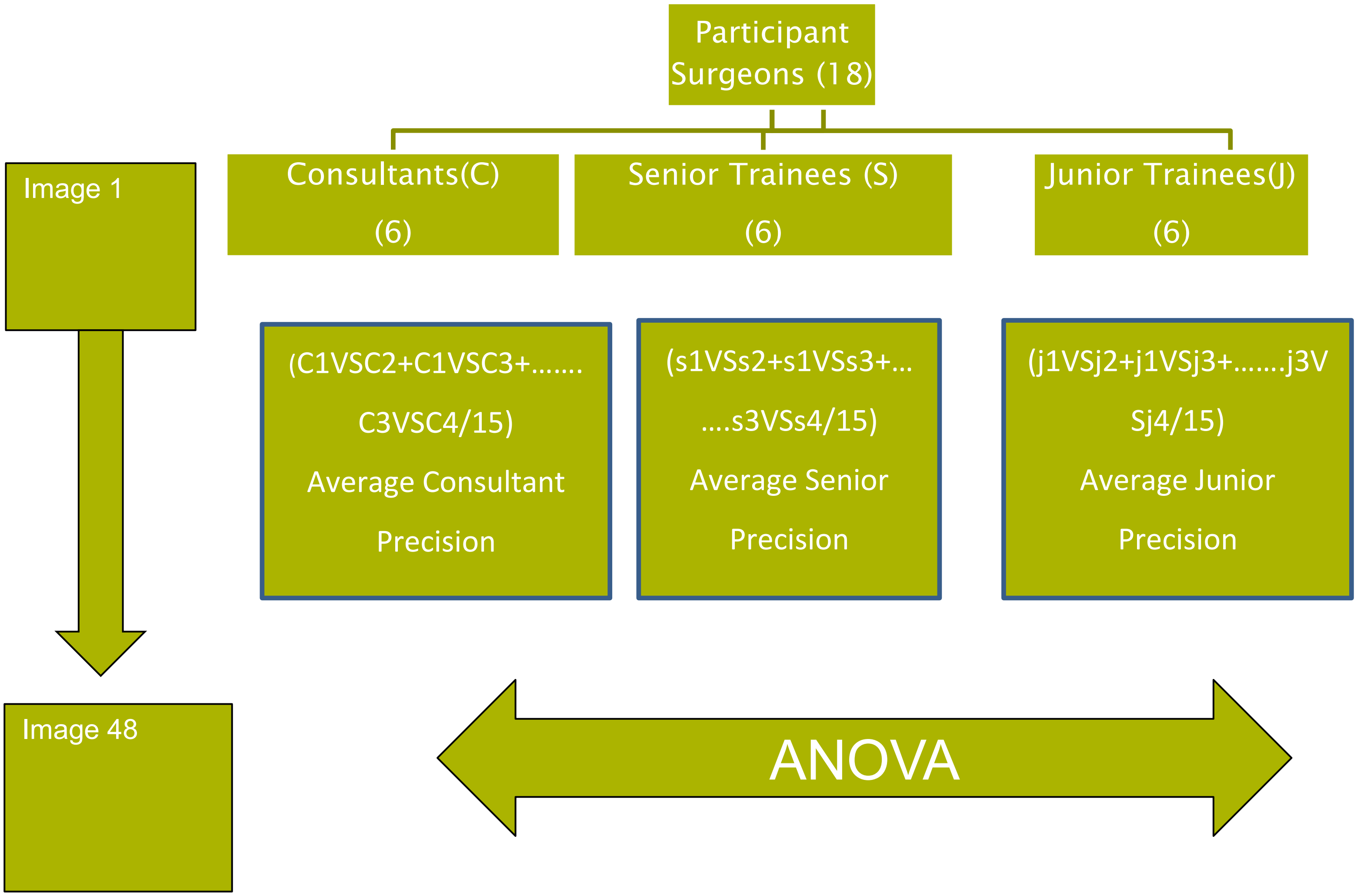


Figure 2.3: Summary sketch of the methodology—Precision

2.14 Analysis of accuracy

As part of the analysis, accuracy of the trainees was also looked at. The interesting methodological problem encountered here was that there was no single "correct" line which all the experts had an agreement upon regarding the correct tissue plane. Therefore there is lack of a 'gold standard' of reference to compare accuracy of trainees. Still it was presumed that the consultants being the experts at performing this operation should be the most accurate and their lines were used as the standard. Lines drawn by each junior and senior trainee were compared to each line drawn by consultants to assess which lines were nearer to the consultants'. These were pairwise comparisons between each junior and consultant and each senior and consultant (Figure 2.4). If the hypothesis is valid, senior trainees who are better at performing these operations than juniors should have their lines being more accurate than juniors, thus adding to the validity of this assessment tool.

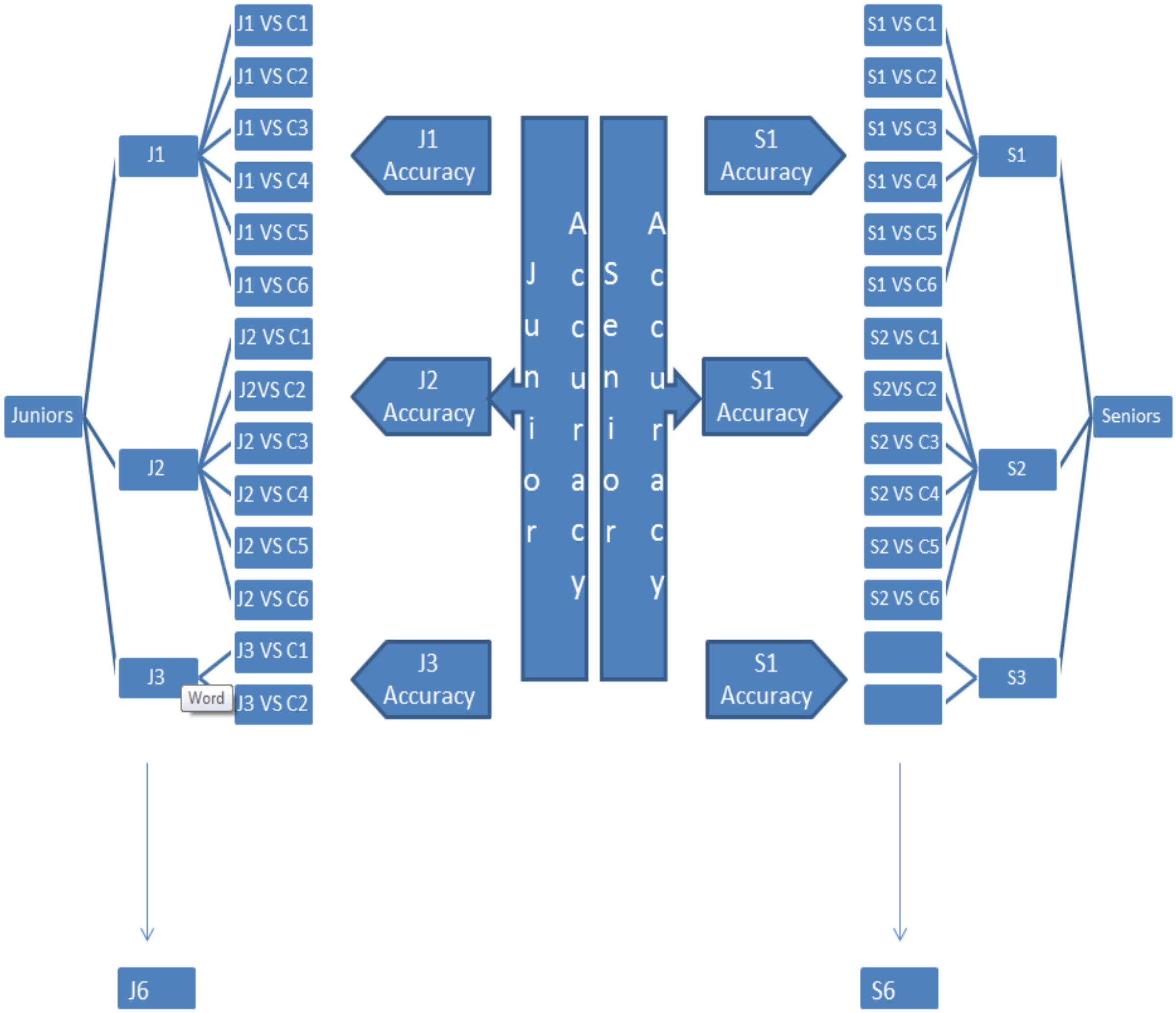


Figure 2.4: **Summary sketch of the methodology used for assessing accuracy of trainees.**

For each image, an individual trainee's lines (J1-J6 and S1-S6) were compared to all consultant lines (C1-C6).

2.15 Sub Analysis

It was noticed that consultants tended to draw longer lines during the course of the study as compared to the trainees. In order to show this, lines drawn by each participant was measured. This was done by the methodology discussed previously but the sum of the point to point distance between each of the nine points was also taken to estimate the variable curvature of the lines. For each participant line A, this total distance of the line i.e. line length for a given image (i) was calculated as the summed distance (d), in pixels, between sequential points on the same line (n1-n8).

$$d(An1 - n8{:}\, i) = \sum_{j=1}^{8} \sqrt{(Anx_j - An'x_j)^2 + (Any_j - An'y_j)^2}$$

Where n'= n+1

This method to calculate for line length measures the line in curvilinear distances thus providing a true measure of the line's length since most of the lines drawn by surgeons are curved rather than straight lines.

2.16 Time record and experience

Although the subjects were not given a specific time to complete the entire exercise, the time required by each participant to complete the exercise was recorded. The subjects were not aware that they were being timed. In addition, the experience which each candidate had with laparoscopic surgery was also recorded.

All the statistical analysis was performed by the SPSS software.

2.17 References

1) Anthony G. Gallagher, Gerald C. O'Sullivan. Fundamentals of Surgical Simulation: Principles and Practice. London, UK: Springer; 2012. p 374.

2) A Novel Method for Assessing Visual Perception of Surgical Planes C.M. Schlachta , Syed Ali, Hammood Ahmed , Roy Eagleson. Manuscript submitted to Surgical Endoscopy.

3) Hausdorff F, Dimension und äusseres Mass, Math Ann. 1918; 79(1-2): 157–79

4) Godfried Toussain, Course CS 507 on Computational Geometry, McGill University, Montreal, QC. 1998. (Date Cited 02/04/2014) Available from: http://cgm.cs.mcgill.ca/~godfried/teaching/cg-projects/98/normand/main.html

5) Panagiota Spyridonos, Georgios Gaitanis, Ioannis D Bassukas, Margaret Tzaphlidou. Gray Hausdorff distance measure for medical image comparison in dermatology: Evaluation of treatment effectiveness by image similarity. Skin Research and Technology. 10/2012; DOI:10.1111/srt.12001

6) Carol E.H. Scott-Conner. The SAGES Manual: Fundamentals of Laparoscopy, Thoracoscopy and GI Endoscopy. Second Edition. New York NY: Springer; 2008.

7) TP Grantcharov, Teaching procedural skills. BMJ. 2008; 336 doi:http://dx.doi.org/10.1136/bmj.39517.686956.47

8) Agachan F, Joo JS, Sher M et al. Laparoscopic colorectal surgery. Do we get faster? Surg Endosc. 1997; 11: 331–335.

9) Senagore AJ, Luchtefeld MA, Mackeigan JM. What is the learning curve for laparoscopic colectomy? Am Surg. 1995; 61: 681–685.

10) Simons AJ, Anthone GJ, Ortega AE *et al.* Laparoscopic-assisted colectomy learning curve. Dis Colon Rectum. 1995; 38: 600–603.

11) Senagore AJ, Duepree HJ, Delaney CP *et al.* Results of a standardized technique and postoperative care plan for laparoscopic sigmoid colectomy: a 30-month experience. Dis Colon Rectum. 2003; 46: 503–509.

3 Integrated Article: Validation of A Novel Method for Assessing Visual Perception of Surgical Planes

Syed Ali MD, Roy Eagleson PhD, Christopher M. Schlachta MDCM

3.1 Introduction

It is considered a fundamental principle of good surgical technique to respect tissue planes during surgery. Tissues planes tend to be avascular and therefore bleeding can be reduced. In addition, there is growing evidence of improved oncologic outcomes associated with adherence to dissection along tissue planes. This has been demonstrated clearly for rectal cancer resections (1). There is compelling evidence to suggest this is also true for colon cancer surgery (2) and hepatobiliary surgery (3).Various academic surgical societies have mandated that trainees have a clear understanding of the correct tissue plane as a prerequisite for competency in the operating room (4).

One of the challenges encountered in teaching trainees to operate within tissue planes is to facilitate the trainee’s recognition of the plane. What is intuitively obvious to the expert surgeon is not always obvious to the trainee. Currently it is believed that this skill is learned though repeated exposure through the course of clinical apprentice-based training. In the current era of surgical training there has been an overall restriction in trainee work hours and thus an overall decrease in trainee exposure in the operating room because of various reasons including patient and trainee safety. Simulation based surgical assessment and training is an attempt to compensate for this and this has been proven to be an effective way to assess and teach surgical trainees (5). There have been numerous successful and validated attempts to assess and teach surgical trainees based on simulation outside the operating room (6, 7). While most simulation based tools focus on assessing and teaching various different aspects of surgical skills, to the authors' knowledge, there have been no attempts at developing and validating a simulation based teaching tool which specifically focuses on trainee's identification of correct surgical tissue planes.

The authors have developed a novel assessment tool to assess this ability. The pilot study conducted on this assessment tool has shown promising results by distinguishing surgeons by their ability to identify surgical planes and this ability correlated strongly with surgeons' level of experience. Thus more experienced surgeons performed better when assessed (8).

The purpose of the current study is to refine and validate this assessment tool by applying it on a larger surgical population. The hypothesis of this study is that this assessment tool accurately distinguishes surgeons based on their ability to identify surgical tissue planes and is able to discriminate between novice and expert surgeons.

3.2 Methods

3.2.1 Selection of Images

A total of 1126 images were initially still captured on a standard video playing computer software, VLC player. These were collected from a video archive consisting of 12 cases of laparoscopic right colectomy. All these surgeries were performed by a single surgeon during 2006-2010 at London Health Sciences Center. These images were thought to represent moments during the surgery where the tissue plane of dissection was visible.

An expert panel consisting of the investigators of this study then reviewed these images. These individuals represent expertise in advanced laparoscopic surgery training, and specific cognitive and perceptual-motor skills. The panel shortlisted a total of 48 images based on the following criteria:

1) The image represents an operative moment where a tissue plane of dissection

should be identified.

2) The ability to identify the dissection plane in these images would likely be of variable difficulty for novices and experts.

3) In selecting an image all three member of the panel had to agree on (1) and (2).

3.2.2 Identification of Subjects

Volunteer participants were recruited from surgical colleagues and trainees in the Division of General Surgery at Western University and the University of Toronto. Participants were divided into three groups based on their level of experience, with a total of six surgeons per group.

1. The consultant group (C) included general surgeons at one of the hospitals affiliated with Western University or the University of Toronto with Royal College certification in general surgery that perform colorectal surgery as a part of their clinical practice. This was considered the expert group with which trainees were referenced. To be included in the consultant group a consultant had to have performed more than 50 laparoscopic right hemicolectomies independently. This number is expected to be associated with reasonable technical competence in performing this procedure. (9, 10, 11, 12)

2. The senior trainees group (S) included residents in their third to fifth years of General Surgery residency at Western University.

3. The junior trainee (J) group included residents in their first two years of general surgery residency at Western University, which is a Royal College accredited Canadian general surgery residency program.

3.2.3 Plane Identification exercise

Images were transferred to an iPad 2 (Apple Computers) and presented in Sketchbook Pro software. Subjects were able to draw on the presented images using a stylus for capacitive touch screens (Slim stylus, Targus, Anaheim, California). For each of the images the subjects were asked to draw a line clearly and precisely where they saw the tissue plane of dissection. Each subjects' line was saved in a separate layer for later analysis.

3.2.4 Analysis of performance

Once data collection was completed, the lines were analyzed by a novel methodology (Figure 3.1). This methodology involved using a distance metric similar to the Hausdorff measure used in computational geometry (13). This distance metric measures how far two subsets of a metric space are from each other. Instead of just taking the shortest distance between two geometric figures, lines or arcs, it also takes into account the maximum distance possible between two objects. This distance metric is different from the Hausdorff distance in that it preserves the sequence of pairs of each line in performing a pairwise Euclidean distance calculation. Each line is sampled into a discrete set of points (same number of points for each line) and the distance metric is established by summing the pairwise distances.

Using MS Paint (Microsoft Office Professional Edition, 2011, Microsoft Corporation, Seattle, WA), each line was iteratively bisected three times resulting in eight equidistant spaces which are represented by nine equidistant points (Figure 3.1). Cartesian coordinates (X,Y) are plotted on each of these nine points, thus giving a numerical representation of the lines' location. For each study group (G), this distance metric between any two lines (a, b) for a given image (*i*) was calculated as the summed distance (*d*), in pixels, between each of the (p=9) coordinate pairs (j) as follows:

$$d(Gab{:}\,i) = \sum_{j=1}^{p} \sqrt{(Gax_j - Gbx_j)^2 + \ (Gay_j - Gby_j)^2}$$

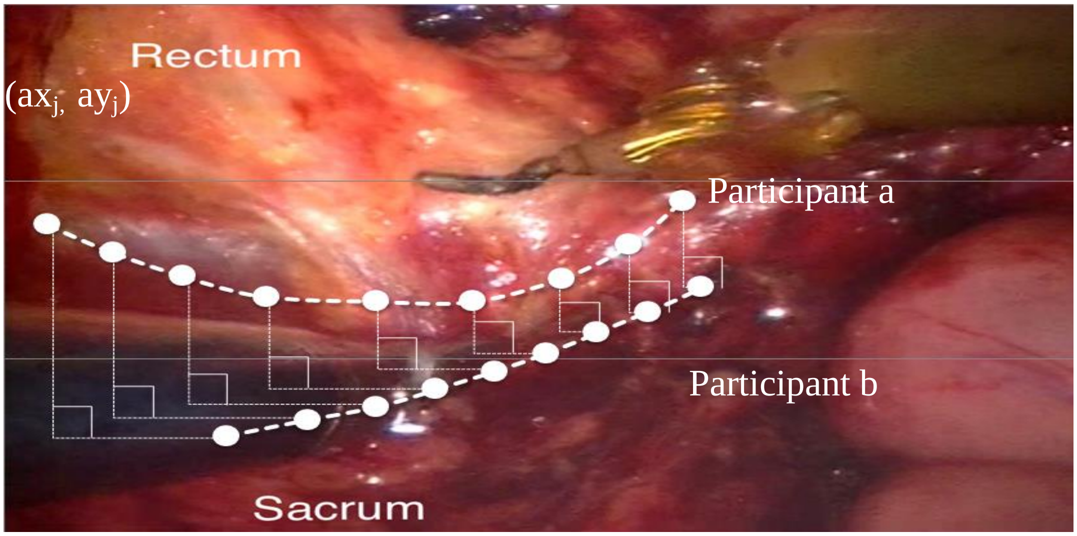


Figure 3.1: Comparing two lines to determine precision.

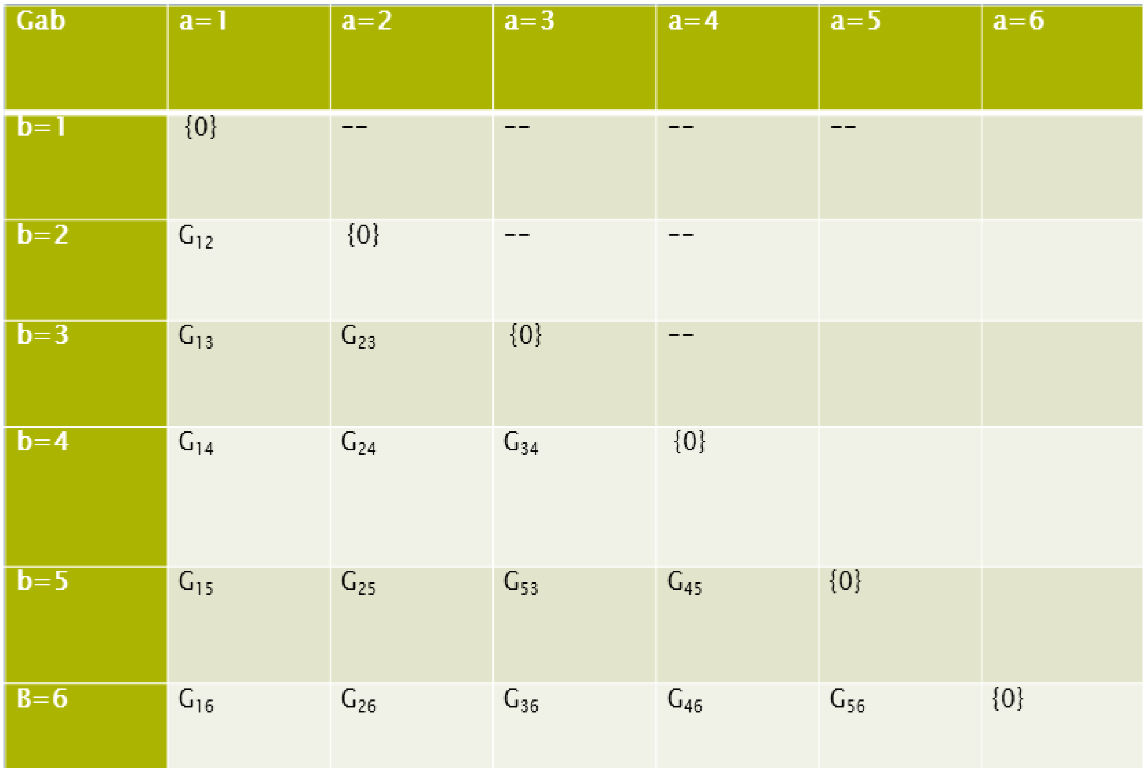
Each line is sampled into a discrete set of points (same number of points for each line) and the distance metric is established by summing the pairwise distances.

.

3.2.5 Analysis of precision:

The lines within a group were compared amongst each other for each image to get intragroup precision values. The smaller the values of the distance metric, the closer the lines are to each other within that group, and the more precise the group. Since there are six participants within a group and each member of a group is being compared with every other member of the same group, there are 15 comparisons per group per image (Figure 3.2).

Mean of these 15 values was calculated to get an estimate of the precision of the particular group for each image. Precision values of each group for each image were compared with the other two groups using one way analysis of variance (ANOVA) as well as student’s t-tests (2 tailed unequal variance). One way ANOVA was used since three groups are being compared amongst each other. Once a statistically significant value was obtained with one way ANOVA between the three groups (F-critical >3.354), pairwise comparison was performed between the groups with 2 tailed students t-test of unequal variance. A Sidak’s correction is used for comparing multiple groups (p<0.017). These steps are summarized in Figure 3.3



Gab	a=1	a=2	a=3	a=4	a=5	a=6
b=1	{0}	--	--	--	--	
b=2	G_{12}	{0}	--	--		
b=3	G_{13}	G_{23}	{0}	--		
b=4	G_{14}	G_{24}	G_{34}	{0}		
b=5	G_{15}	G_{25}	G_{53}	G_{45}	{0}	
B=6	G_{16}	G_{26}	G_{36}	G_{46}	G_{56}	{0}

Figure 3.2: Obtaining the number of precision comparisons per group.

When six participants are compared to each other, 15 comparisons of precision per group are obtained.

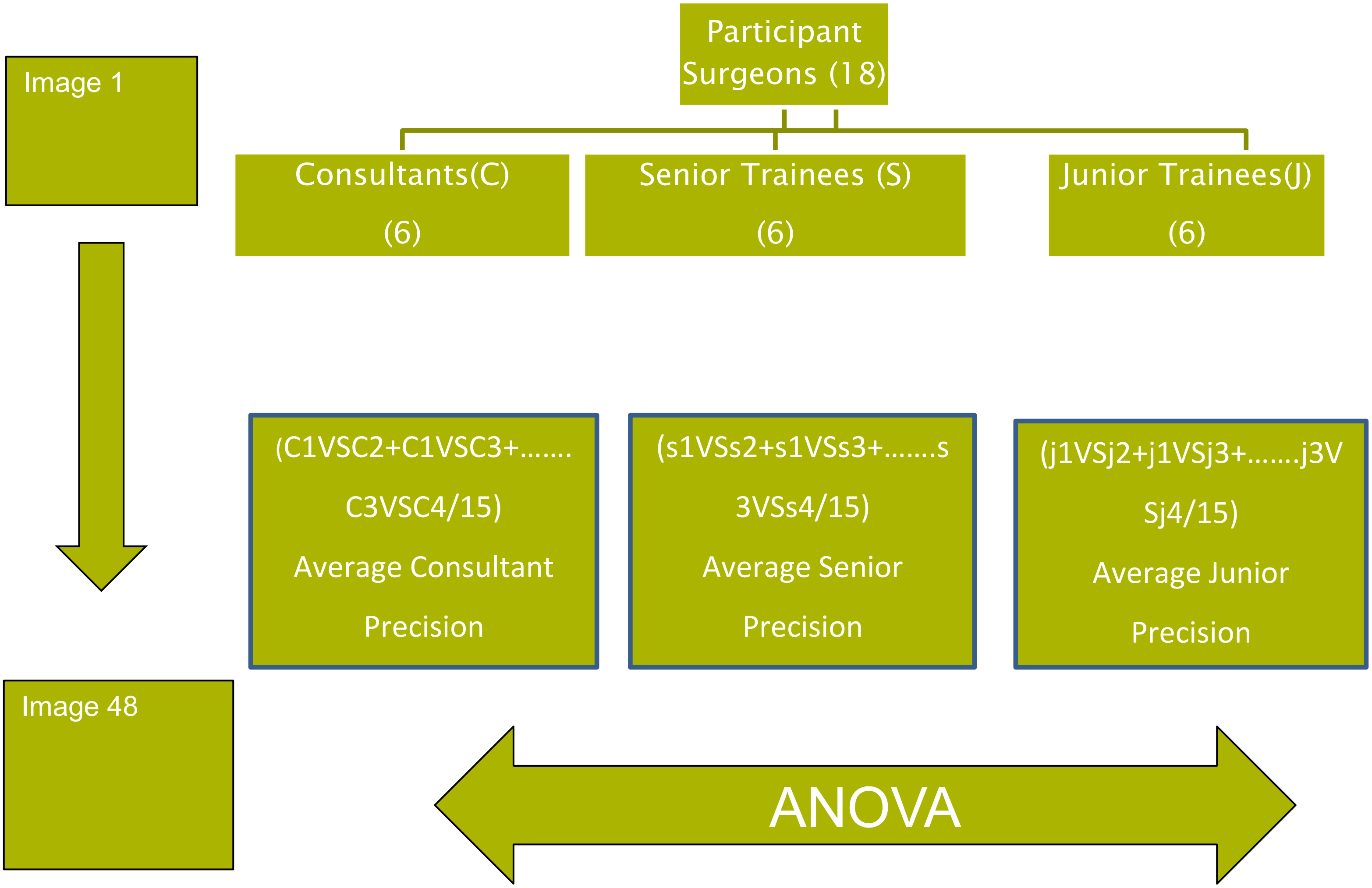


Figure 3.3: Summary of the Methodology-Precision

3.2.6 Analysis of accuracy

As part of the analysis accuracy was also looked at. The interesting phenomena encountered here was that there was no one line which all the experts had an agreement upon regarding the correct tissue plane. Therefore there is no standard of reference to compare accuracy of trainees. Still it was felt that the consultants being the experts at performing this operation should be the most precise and their lines were used as the standard. Lines drawn by each junior and senior trainee were compared to each line drawn by consultants to assess which lines were nearer to the consultants. These were pairwise comparisons between each junior and consultant and each senior and consultant (Figure 3.4).

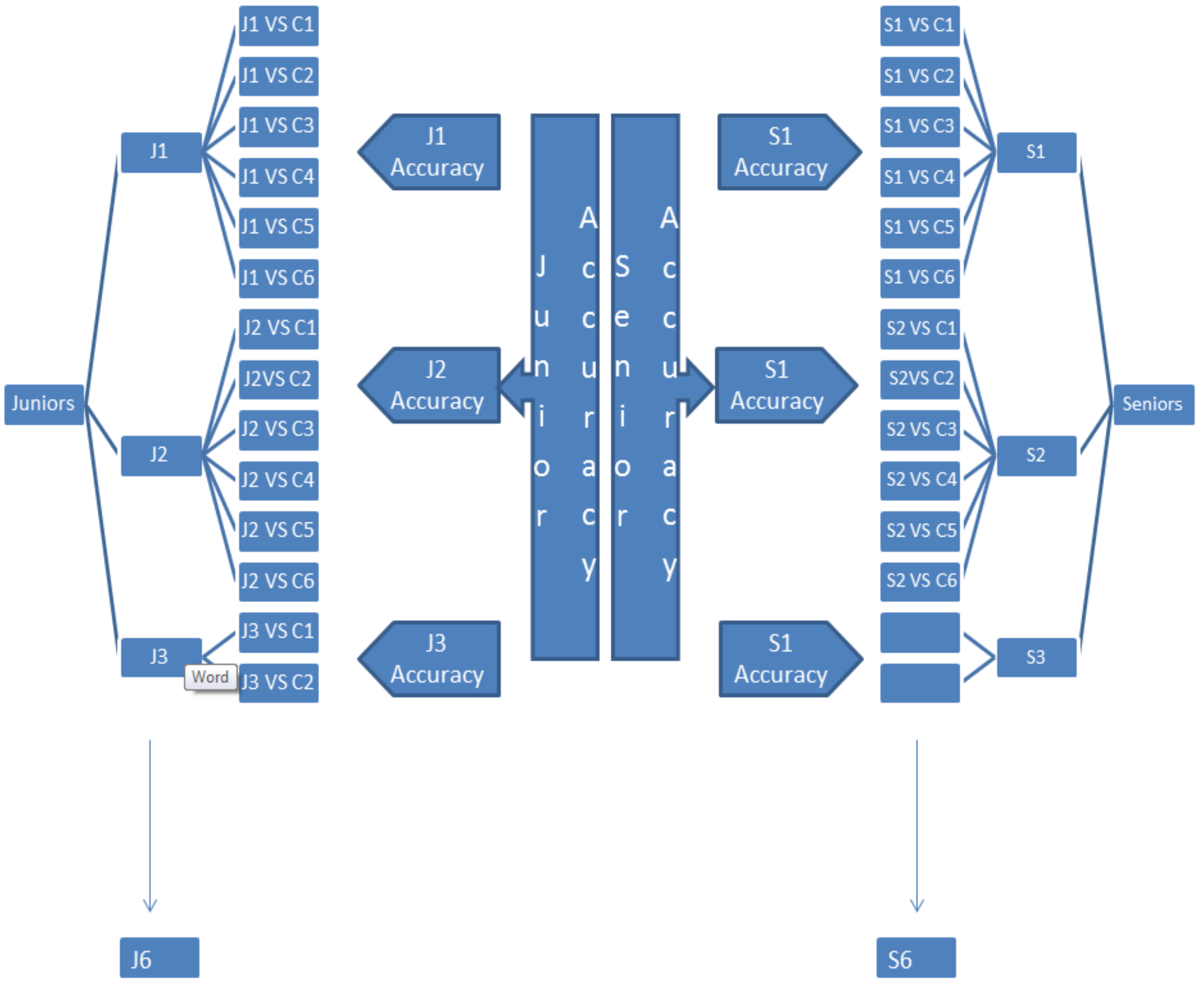


Figure 3.4: Summary sketch of the methodology used for assessing accuracy of trainees.

For each image, an individual trainee's lines (J1-J6 and S1-S6) were compared to all consultant lines (C1-C6).

3.2.7 Sub Analysis

Consultants tend to draw longer lines during the course of the study as compared to the trainees. In order to show this, lines drawn by each participant was measured. This was done by the methodology as discussed previously but the sum of the point to point distance between each of the nine points was also taken to estimate the variable curvature of the lines. For each participant line A, this total distance of the line i.e. line length for a given image (*i*) was calculated as the summed distance (*d*), in pixels, between sequential points on the same line (n1-n8).

$$d(An1 - n8{:}\, i) = \sum_{j=1}^{8} \sqrt{(Anx_j - An'x_j)^2 + \ (Any_j - An'y_j)^2}$$

Where n' = n+1

This method to calculate for line length measures the line in curvilinear distances thus providing a true measure of the lines length since most of the lines drawn by surgeons are curved rather than straight lines.

3.2.8 Time Record and Experience

Time required by each participant to complete the exercise was recorded, as was each candidate's self-reported experience with laparoscopic surgery. Subjects were not given a specific time to complete the entire exercise and they were not aware that they were being timed. In addition, the experience which each candidate had with right hemicolectomies was also recorded.

The statistical analysis was performed by the SPSS software.

3.3 Results

A total of 18 participants in three groups were tested on a total of 48 images. Analysis was performed as per the proposed methodology on a total of 864 drawn lines (18 participant lines × 48 images).

3.3.1 Analysis of precision

Out of all 48 images, one way analysis of variance (1-WAY ANOVA) revealed statistically significant differences amongst the three groups ($P<0.05$, $F>F_{crtical=}3.21994$) in 25 out of 48 images. On pairwise comparison, consultants were significantly more precise than trainees (Seniors, Juniors or both) in 14 images. Seniors were significantly more precise than Juniors in nine images. Table 3.1 highlights this data with statistically significant results. Figure 3.5 is a flow chart summarising the precision results and figure 3.6 describes the distribution of all the statistically significant pairwise comparisons.

Seniors were significantly more precise than consultants in five images. On direct review of these images it was appreciated that in three cases trainees, albeit more precise, were drawing their lines in a different location on the image. They had the wrong tissue plane. In a further two images, there were significant differences of opinion between consultants as to the location of the ideal tissue plane resulting in orthogonally oriented lines. This seems to generate much larger measures on the modified Hausdorff analysis.

3.3.2 Analysis of accuracy

The differences in accuracy between the senior and junior trainees reached statistical significance in favor of seniors in 14 images whereas junior trainees were statistically more accurate in eight images (Table 3.2 and Figure 3.7).

3.3.3 Duration per participant

Each participant’s time required to complete the exercise is recorded in Table 3.3. The average time taken by Consultants was the greatest amongst the three groups. Average Seniors’ time was shorter than Consultants’ with Juniors’ time being the shortest. These differences reached statistical significance on one-way ANOVA and remained significant on pairwise comparison with a two-sided student’s t- test between Consultants and both levels of trainees (C vs J and C vs S). The time taken by Junior and Senior trainees (J vs S) was however not significantly different (Table 3.3).

3.3.4 Accuracy of each trainee group

The senior trainees scored better than the junior trainees as a group. When the group average is compared between seniors and juniors, these differences do not reach statistical significance.

There was considerable variability in the self-reported experience with laparoscopic and open right hemicolectomies done by trainees (Tables 3.4 A and B).

3.3.5 Length of lines

Each participant's lines drawn on each image were also analysed by the modified Hausdorff distance calculation formula. The average line length of consultants was statistically longer as compared to lines drawn by trainees (Tables 3.5 A and B).

Table 3.1 : Selected images showing comparison of precision between consultant and trainee groups.

Image**	Mean Distance±S.D in pixels			ANOVA(P)*	C vs S (P)	C vs J(P)	J vs S(P)
	Consultant	Senior	Junior				
7	445±221	1778 ±771	1710±707	1.93E-07*	5.55E-06*	4.2E-06*	0.800914
11	597±271	1155±657	1470±694	0.0006*	0.007*	0.0002*	0.212
42	853± 496	2080±1061	1532± 551	0.0002*	0.0006*	0.0006*	0.09

46	853±309	1382±723	1243±611	0.04*	0.01*	0.01*	0.57
22	932±514	1907±731	1770±761	0.0005*	0.0003*	0.001*	0.62
5	969±398	1848±627	1105±335	0.00001*	0.0001*	0.3	0.0007*
23	973±442	1614±927	1722±1089	0.047*	0.02	0.02	0.77
8	1009±409	1560±700	2111±1259	0.0056*	0.016*	0.0055*	0.152952
2	1056±402	2021±1139	1254±616	0.003*	0.006*	0.3	0.03
17	1152±351	2137±807	1484±520	0.0001*	0.0004*	0.05	0.015*
6	1197± 384	1996± 661	2014± 787	0.001*	0.005*	0.001*	0.34
27	1497±669	2114±686	2344±736	0.005*	0.01*	0.01*	0.38
19	1688±830	2500±1201	2732±1081	0.02*	0.04	0.006*	0.58
39	1282±822	2143±669	1178±602	0.0008*	0.004*	0.691	0.0002*
14	1364±574	952±511	865±584	0.04*	0.047	0.026	0.667343
41	989±446	847±441	1390±301	0.002*	0.36	0.01*	0.0007*
9	1072±563	608±210	1973±945	4.707E-06*	0.008*	0.00427*	0.00006*
25	1115±535	829±350	1290±523	0.04*	0.09	0.09	0.009*
47	1205±754	782±228	1450±660	0.01*	0.05	0.05	0.001*
13	1270± 591	770±294	1743±774	0.001*	0.008*	0.121	0.002*
20	1680± 935	996±507	2183± 1158	0.005*	0.01*	0.2	0.001*
18	1684 ±1074	1297± 1455	3833±753	2.88E-07*	0.264	0.00009*	5.522E-06*

40	1783± 733	875±279	1925±748	5.77E-05*	0.002*	0.61	0.0000785*
36	1967±1158	1121±386	1497±674	0.02*	0.01*	0.1	0.07
26	2191±801	1566±412	2489±1222	0.02*	0.01*	0.43	0.01*

*statistically significant difference with Sidak's correction for multiple comparisons ($P < 0.017$)

** Images are arranged in descending order of consultant precision.

SD= Standard Deviation

Table 3.2: Images showing comparison of accuracy between trainee groups.

Image	More accurate	Mean distance in pixels ±SD		P value(Accuracy)
	S or J	S vs C	J vs C	
1	S	1237±770	1564±817	0.08
2	J	1512±903	1171±517	0.03*
3	S	1793±815	2503±911	0.0008*
4	J	1502±711	1359±629	0.47
5	J	1503±729	1106±450	0.0007*
6	S	1552±663	1698±748	0.38
7	S	1218±883	1390±774	0.34

8	S	1400±897	2229±873	0.001*
9	S	866±329	1453±686	0.00003*
10	S	1511±823	1822±852	0.11
11	S	1083±483	1196±542	0.35
12	J	1123±763	956±473	0.27
13	S	1019±503	1467±767	0.004*
14	J	1422±494	1360±581	0.6
15	S	721±300	1128±567	0.0003*
16	S	1428±906	1597±893	0.4
17	J	1712±702	1518±482	0.17
18	S	1699±1218	3148±1258	4.65E-06*
19	S	2250±1001	2260±1226	0.12
20	J	2099±1200	1861±1139	0.39
21	J	1251±545	1159±509	0.46
22	S	1385±668	1540±897	0.42
23	S	1495±915	1509±681	0.94
24	S	1113±552	1453±808	0.04*

25	S	925±424	1414±685	0.003*
26	S	1779±735	2330±903	0.005*
27	J	1923±650	2006±664	0.5
28	J	1745±980	1285±570	0.01*
29	J	2260±1189	2550±1060	0.27
30	S	1969±800	1807±848	0.4
31	S	2117±1287	2482±1158	0.2
32	S	1512±750	2129±1085	0.006*
33	J	2655±803	2179±1121	0.04*
34	S	2870±1509	2389±1387	0.16
35	J	1557±779	1597±816	0.83
36	J	2375±1036	1648±843	0.001*
37	J	1659±704	1697±776	0.82
38	J	1923±930	1491±721	0.03*
39	J	2699±951	1335±783	0.0000000638*
40	S	1552±814	1734±777	7.85E-05*
41	S	925±432	1191±444	0.0007**

42	J	1494±716	1310±590	0.09
43	S	1880±858	2337±838	0.02*
44	J	1317±579	985±474	0.009*
45	J	1055±445	961±468	0.42
46	S	1294±412	1376±765	0.57
47	S	1184±751	1553±1034	0.001*
48	S	1134±598	1165±655	0.99

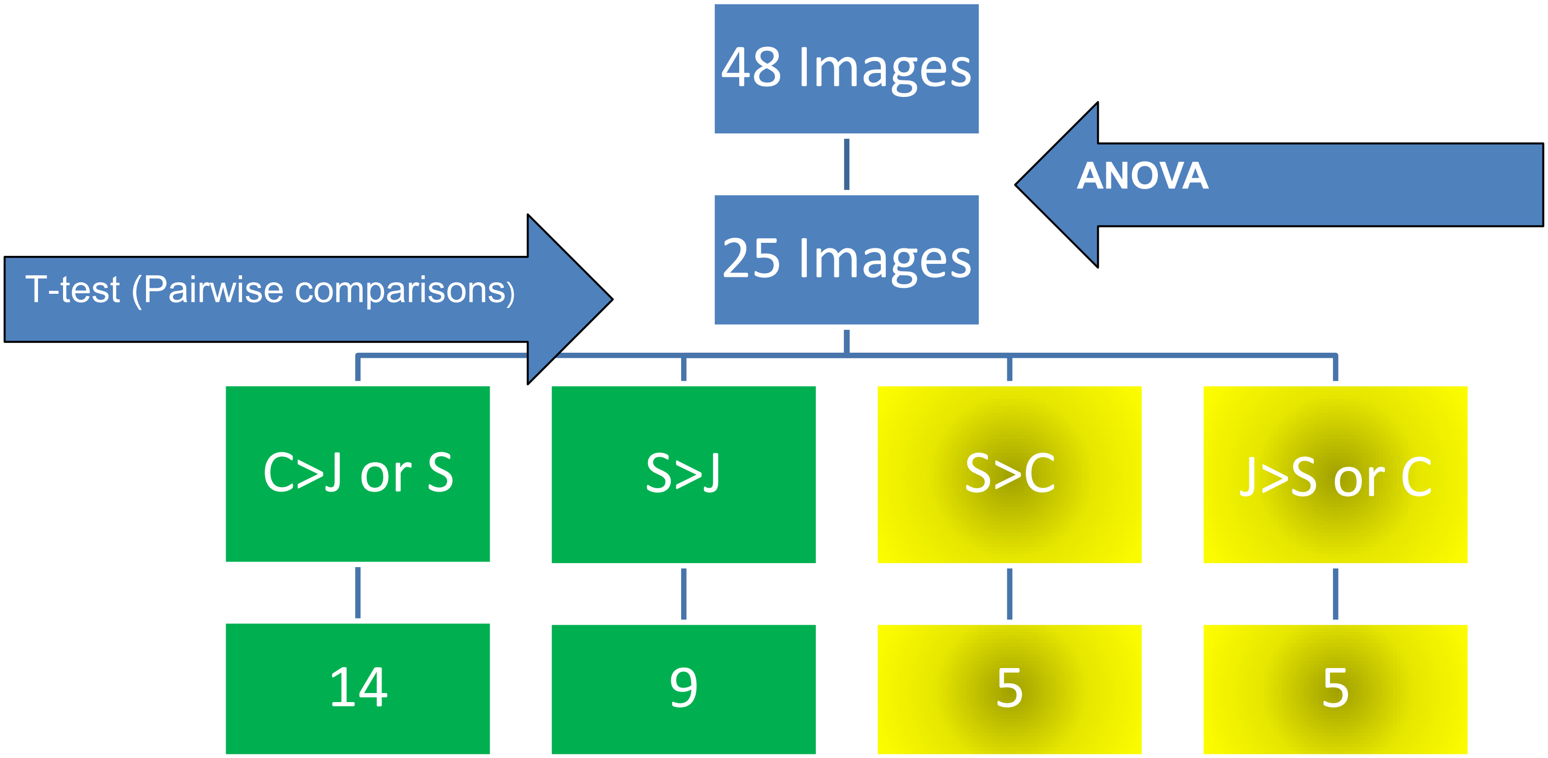


Figure 3.5: Images with statistically significant results based on precision.

The green shaded boxes represent the number of images supporting the hypothesis while the yellow shaded boxes indicate images with conflicting results.

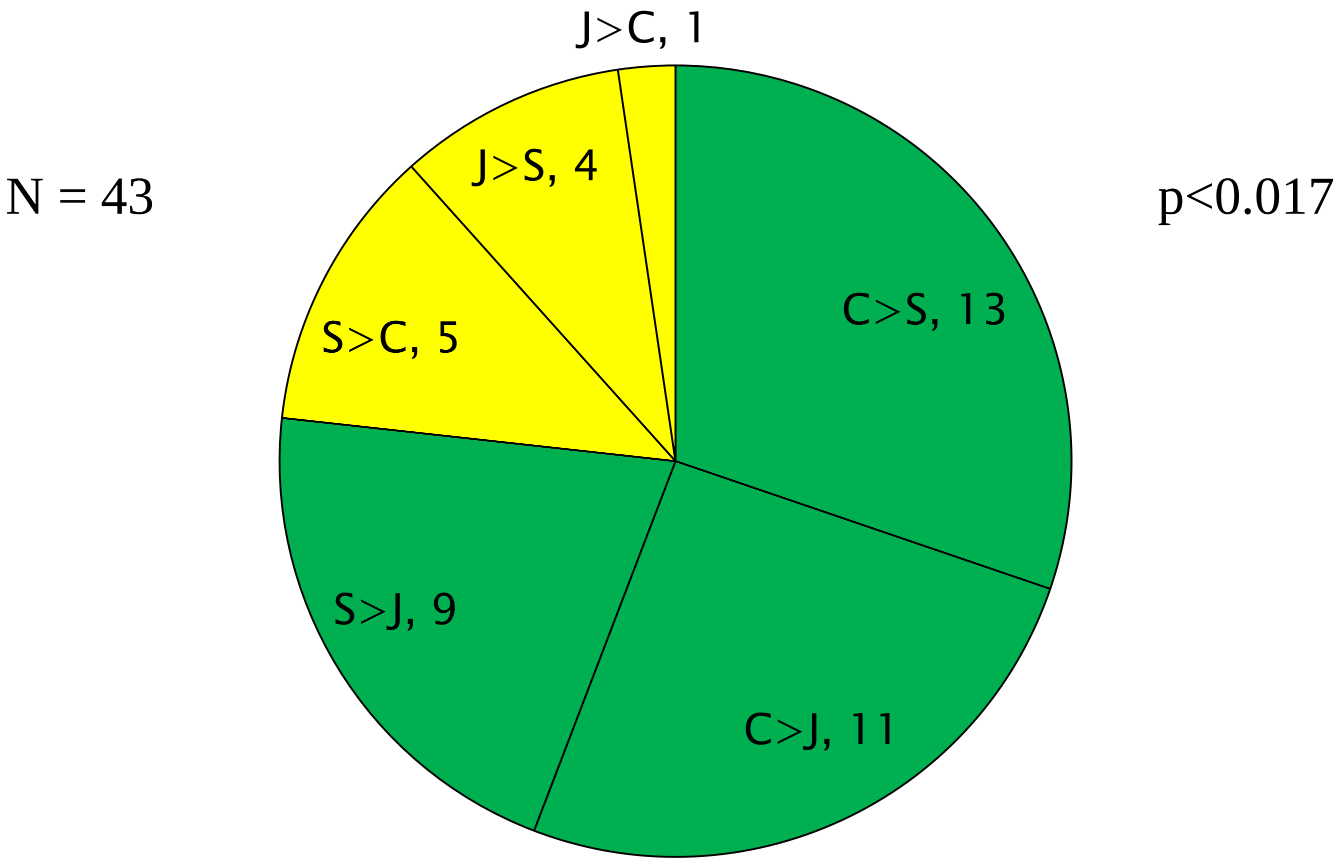


Figure 3.6: Distribution of all the statistically significant pairwise comparisons amongst the three groups.

When the Sidak's corrected P value is used for multiple comparisons (P<0.017), 43 pairwise comparisons show statistical significance. Three quarters of the results support the hypothesis (indicated by the green area) while one quarter consists of conflicting results (yellow area).

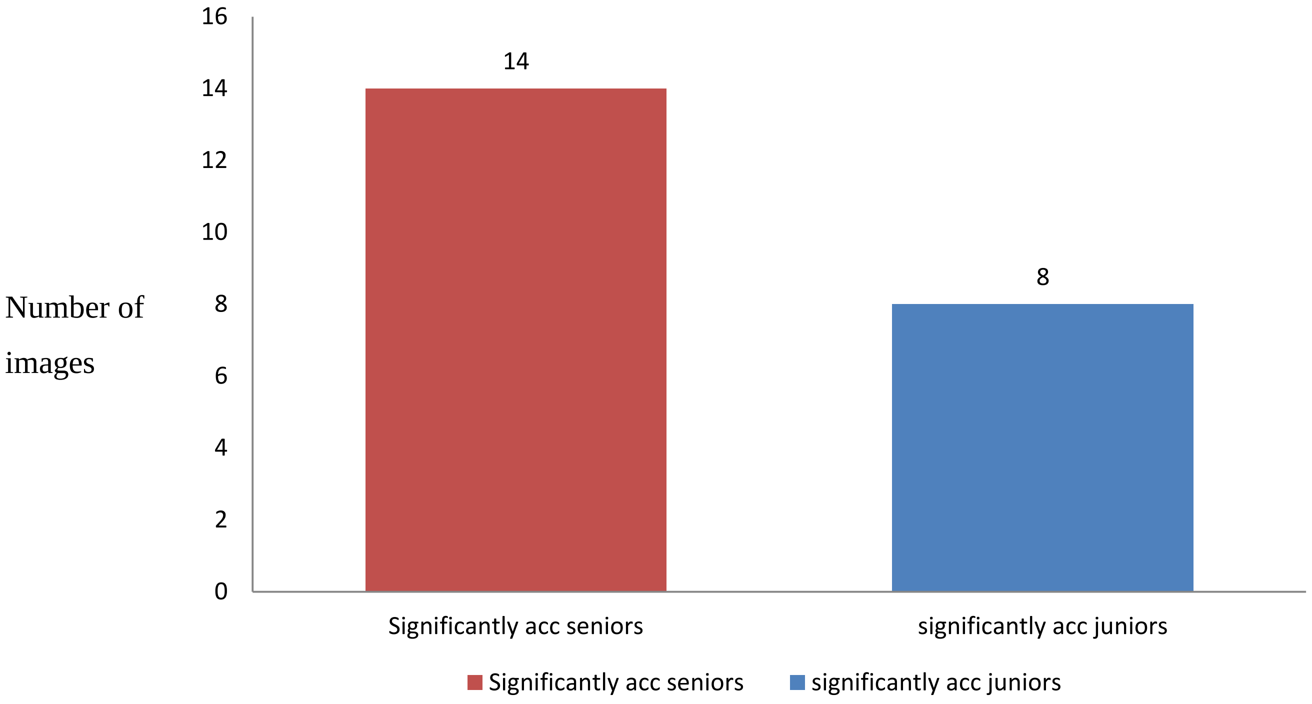


Figure 3.7: Bar chart representing accuracy of trainees.

The y-axis indicates the number of images in which the respective trainee group were more accurate.

Table 3.3: Total time taken by each participant.

Consultants	Time taken	Seniors	Time Taken	Juniors	Time Taken
C1	55	S1	42	J1	29
C2	64	S2	48	J2	32
C3	48	S3	40	J3	31
C4	39	S4	46	J4	28
C5	59	S5	48	J5	36

C6	62	S6	43	J6	35
Mean time (min)±SD	60.75±14.3		40.38±10.7		33±3.54
T-test- C vs S	0.006*	C vs J	0.0009*	J vs S	0.09

Table 3.4 A: Comparing accuracy score of senior trainees to their experience.

Seniors	S1	S2	S3	S4	S5	S6	Mean±SD
Experience	40	30	10	3	2	5	15±16.04
Accuracy	1689.621	1650.319	1313.231	1518.531	1634.828	1658.104	1577±131.3

Table 3.4 B: Comparing accuracy score of junior trainees to their experience.

Junior Trainees	J1	J2	J3	J4	J5	J6	Mean±SD
Experience- number of cases	3	3	15	20	9	10	10±6.69
Accuracy	1703.111	1646	1807	1487	1603	1682	1655±106.91

Table 3.5A: Length of lines drawn by participants for each image.

Image	senior lines	Consultants	Juniors	Image	Senior Line Length	C-line Length	J-line Length
1	185.5787	240.9684	251.9261	25	364.5669	500.452	372.0759
2	239.086	356.7852	323.3773	26	332.3228	517.3985	375.1137
3	304.664	523.4079	350.7598	27	341.8356	576.6979	422.4217
4	427.6495	582.3615	466.2588	28	338.4264	515.9522	399.4577
5	431.9403	316.2208	306.0027	29	428.386	460.8107	569.4562
6	361.1794	365.2964	340.6466	30	311.3882	591.1223	370.2004
7	385.5398	331.1948	406.2185	31	328.6989	611.3787	480.4163
8	309.1987	331.0611	409.9014	32	379.6094	393.9628	351.7307
9	368.001	462.5503	403.3229	33	341.3999	584.9601	411.7855
10	432	533	487	34	265.5962	465.0855	394.1451
11	278.3433	442.06	262.6241	35	353.39	432.234	388.5816
12	296.12	363.6632	399.64	36	260.0219	491.7312	394.854
13	450.0211	621.224	463.6962	37	280.9558	456.7309	336.1078
14	409.0713	702.437	428.2759	38	346.1049	428.6113	365.1958

15	410.3743	437.088	512.389	39	379.7473	700.1542	453.3542
16	355.0476	569.69	327.4889	40	225.0254	421.7729	361.6587
17	320.9317	548.3923	295.9106	41	184.951	209.1826	278.2078
18	244.3948	407.85	324.7139	42	274.115	362.75	368.6424
19	391.5561	491.085	302.9425	43	314.4082	381.2255	346.0547
20	238.431	316.25	262.4664	44	207.1525	298.6695	237.7277
21	467.6933	567.203	391.3248	45	251.2541	371.6456	323.5073
22	439.4454	458.83	281.3172	46	256.2958	484.12	417.723
23	303.4169	434.87	359.7755	47	356.3475	423.9026	438.0473
24	341.2639	485.981	394.5471	48	244.0614	348.783	263.3926

Table 3.5 B: Comparing average line length drawn by consultants, seniors and juniors.

Consultants ± SD	Seniors ±SD	Juniors ±SD
456.6418 ± 72	372.3413 ± 109	328.271 ± 71
Pairwise t-test for line length		
C versus S	C versus J	J versus S
1.53174E-12	1.16679E-07	4.05259E-05

3.4 Discussion

Based on these results it can be fairly concluded that the assessment tool designed to measure the ability to identify tissue planes is successful in doing so. It is able to distinguish surgeons based on their levels of experience on measures of accuracy and precision. When all the images are considered together, the senior most participants (C) appear to be the most precise while the junior most (J) are the least precise. Similarly senior trainees are more accurate than juniors trainees when compared with consultants.

This study is a robust attempt to validate the pilot study. The hypothesis was supported by the results on statistical grounds by using an ANOVA with traditional critical F values and a two sided t test for pairwise comparisons. It is known from previously done studies that accuracy and precision based analysis are a valid way to assess surgical performance (17). Employing an expert panel to select images added to the content validity of the assessment tool. Having access to a large archive of images gave investigators considerable extent of leverage in selecting the final set of images (48 out of 1000+ images). Though the sample size was still not very large, one more participant per group was recruited then what was required by the power analysis. A heterogeneous sample of surgeons with variable experience and training was utilized, which is fairly representative of the North American General Surgical population.

The visual processing of the spatial relations of image properties is known as visuospatial ability. The visuospatial ability of a surgeon plays a very important role in identifying tissue planes. There has been sufficient evidence in literature pointing towards the importance of visuospatial abilities in surgical task performance (18 and 19). Literature review shows that experience results in enhancement of performance on visuospatial tasks as pointed out by Keehner *et al*. (18). These finding can also be appreciated in this study with consultants performing the best in terms of precision.

Current training in surgery is going through drastic changes with restrictions in work hours because of patient and trainee safety issues (20). This has resulted in reduction in the amount of exposure the trainee gets in the operating room. Simulation can play a very useful role in narrowing this gap and is on the top agenda of most North American surgical education and accreditation societies (21). Current simulation models focus on overall performance and the assessment metrics employed by these simulations are not very content specific (22, 23). Additionally, literature pertaining to visual perception suggests that it is complex heterogeneous skill with various subdomains. In order to understand its role in surgical performance there is a need to assess a surgeon's ability on specific tasks which require visuospatial ability (24). This study is an attempt towards the development of a meaningful simulation tool which attempts to develop a valid assessment metric for a very content specific surgical skill, an ability which most simulation tools lack (22). The performance of each junior and senior trainee was compared on this assessment tool, thus opening the door towards practical applications of this tool for assessing surgical trainee performances in actual clinical settings.

Interestingly, time taken by the participant seems to be inversely proportional to experience with the consultants taking the most time to complete the exercise. This might be because consultants were putting an extra effort to complete the exercise and since they had the most familiarity with the extent of visual detail in these images, so they spend extra time in visually analyzing the depth of information in these images.

An important question which can be raised regarding the content validity of this tool is that is the surgeon's ability to identify tissue planes being accurately measured? This methodology is based on the assumption that this is an exclusively visuospatial task but it is well known that surgical tasks often fall into overlapping domains of human intelligence, visuospatial and psychomotor skills. In actual clinical settings, surgeons tend to use refined manual techniques to expose the tissue such as to improve their visual stimulus and thus this manual dexterity does tend to play a role in facilitating dissection in the correct tissue plane. The pictures used in this study are images in which the tissue plane is already exposed. It can still be stated with confidence that this tool does assess the visuospatial aspect of the ability required to identify a tissue plane.

Seniors were significantly more precise than consultants in five images because of reasons explained above. The lower consultant precision values seen in some of these images can be addressed in the future by getting more input from consultants when selecting images for testing.

The results indicate the significant line length discrepancy between the groups. This was attributed to the fact that consultants were surer about the location of the tissue plane in the whole image and thus were drawing bigger lines. It is postulated that this line length discrepancy might result in decreasing the variability and accuracy estimate of the modified Hausdorff distance metric. When the line length is smaller, the nine equidistant points might be tightly clustered together and lines having tightly clustered Cartesian points might give smaller differences in the Hausdorff distance. To adjust for line length in the distance metric could be one of the future projects arising from this study.

An interesting finding encountered is the large variability in trainee experience with right hemicolectomy. Some of the junior trainees have more experience with this procedure then the seniors and both groups have large standard deviations. This assessment tool measures a unique surgical ability which is not necessarily acquired specifically by exposure to a right hemicolecotmy. Identifying the surgical tissue plane between the mesocolon and the reteroperitoneum is a skill which is very commonly employed in a large number of abdominal procedures and the skills learned are probably transferable from one procedure to another. Reported experience with right hemicolectomy does not necessarily reflect total operative experience on abdominal procedures. In order to evaluate the impact which trainee experience has on the analysis, the senior and junior trainee groups were re-divided based on the trainee’s specific experience of a right hemicolectomies. The results of this diminished the statistical significance of the difference in the measures obtained for all three groups, thus potentially adding an element of contamination to the data and did not contribute positively towards the analysis. This supports findings from earlier studies indicating that surgical trainees with similar experience might have very high variability in their learning curve and this has been demonstrated specifically for tasks involving visuospatial abilities (17). It would seem that surgical trainees are using spatial reasoning to identify tissue planes rather than just memory of the specific procedure.

Also the accuracy data tends to point towards the junior trainees being more accurate than the seniors in a few images. The differences can be attributed to the variability in trainee learning curve as explained above but can also be because of a lack of a clear gold standard for determining accuracy. The consultants are being used based on their clinical experience but it is clear that all the consultants do not have an absolute agreement on the precise location of the dissection plane.

To conclude, this study is a unique exploratory study which has a tremendous potential in the field of simulation based surgical education. The addition of dynamic visuals, adjusting for line length discrepancy and getting more input from experts are some of the future directions this study might take. It can be concluded that this assessment metric has the potential to be used in actual clinical settings and can play an important role in the assessment and training of surgical trainees.

3.5 References

1. Joint Committee on Surgical Training: Guidelines for the award of a CCT in General Surgery (Internet). JCST, London UK. (Date cited: 02/04/14). Available from: http://www.jcst.org/quality_assurance/Docs/cct_guidelines_gs

2. West NP, Morris EJA, Rotina O, Cairns A, Finan PJ and Quirke P. Pathology grading of colon cancer surgical resection and its association with survival: a retrospective observational study. Lancet Oncol. 2008; 9: 857–65.

3. Imamura H, Seyama Y, Kokudo N, Maema A, Sugawara Y, Sano K, *et al.* One thousand fifty-six hepatectomies without mortality in 8 years. Arch Surg. 2003; 138(11): 1198-1206.

4. Association of program directors of colon and rectal surgery: Global Rating scale for Abdominal Colectomy, Ileostomy and Hartmann Closure (Internet) APDCRS, Taylor, MI. Available from: http://www.apdcrs.org/GlobalAssessmentAll.pdf

5. Dawe SR, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, Maddern GJ.A systematic review of surgical skills transfer after simulation-based training: laparoscopic cholecystectomy and endoscopy. Ann Surg. 2014 Feb;259(2):236-48. doi: 10.1097/SLA.0000000000000245.

6. Martin J, Regehr G, Reznick RK, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br JSurg. 1997;84:273–278.

7. Sidhu, R. S., Tompa, D., Jang, R., Grober, E. D., Johnston, K. W., Reznick, R. K., & Hamstra S. J. Interpretation of three-dimensional structure from two-dimensional endovascular images: implications for educators in vascular surgery. Journal of Vascular Surgery. 2004; 39, 1305-1311

8. A Novel Method for Assessing Visual Perception of Surgical Planes, Christopher M. Schlachta , Syed Ali, Hammood Ahmed, Roy Eagleson. Manuscript submitted to Surgical Endoscopy.

9. Agachan F, Joo JS, Sher M et al. Laparoscopic colorectal surgery. Do we get faster? Surg Endosc. 1997; 11: 331–335.

10. Senagore AJ, Luchtefeld MA, Mackeigan JM. What is the learning curve for laparoscopic colectomy? Am. Surg. 1995; 61: 681–685.

11. Simons AJ, Anthone GJ, Ortega AE et al. Laparoscopic-assisted colectomy learning curve. Dis Colon Rectum 1995; 38: 600–603.

12. Senagore AJ, Duepree HJ, Delaney CP et al. Results of a standardized technique and postoperative care plan for laparoscopic sigmoid colectomy: a 30-month experience. Dis Colon Rectum. 2003; 46: 503–509.

13. Hausdorff F, Dimension und äusseres Mass, Math Ann. 1918; 79(1-2): 157–79

14. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. Stud Health Technol. Inform. 1998; 50:124-130.

15. Darzi A, Mackay S. Assessment of surgical competence. Qual Health Care. 2001; 10:(Suppl 2):ii64–ii69.

16. Risucci D, Cohen JA, Garbus JE, Goldstein M, Cohen MG. The effects of practice and instruction on speed and accuracy during resident acquisition of simulated laparoscopic skills. Curr. Surg. 2001; 58:230-235.

17. Martin J, Regehr G, Reznick RK, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br JSurg. 1997;84:273–278.

18. M. Keehner, Y. Lippa, D.R. Montello, F Tendick, M. Hegarty. Learning a Spatial Skill for Surgery: How the Contributions of Abilities Change With Practice. Appl. Cognit. Psychol. 2006; 20: 487–503. DOI: 10.1002/acp.1198

19. Kyle R Wanzel, Stanley J Hamstra, Dimitri J Anastakis, Edward D Matsumoto, Michael D Cusimano. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet. 2002; 359: 230–231

20. American Academy of Graduate Medical Education: Recommendation of Committee on work hour restrictions (Internet Article). AAGME, Chicago, IL (Date cited 02/04/2012) Available from https://www.acgme.org/acgmeweb/tabid/271/GraduateMedicalEducation/DutyHours.asp

21. Position of the American College of Surgeons on Restrictions of work hours.(Internet).ACS Chicago IL.(Date cited: 02/04/2014) Available from http://www.facs.org/education/statement.pdf

22. Leanne M. Sutherland, Philippa F. Middleton, and Guy J. Maddern,. Surgical Simulation A Systematic Review. Ann Surg. Mar 2006; 243(3): 291–300.doi: 10.1097/01.sla.0000200839.93965.26

23. David A. Cook, Rose Hatala, Ryan Brydges, Benjamin Zendejas, Jason H. Szostek, Amy T. Wang, Patricia J. Erwin, Stanley J. Hamstra. Technology-Enhanced Simulation for Health Professions Education: A Systematic Review and Meta-analysis. JAMA. 2011; 306(9):978-988. doi:10.1001/jama.2011.1234.

24. Dimitri J. Anastakis, Stanley J. Hamstra, Edward D. Matsumoto.Visual-Spatial Abilities in Surgical Training. Am J Surg. 2000; 179:469–471

25. Jan-Maarten Luursema, Willem B. Verwey, Remke Burie. Visuospatial ability factors and performance variables in laparoscopic simulator training. Learning and Individual Differences. 2012; 22: 632–638

4 Results

The results were divided primarily based on precision versus accuracy analysis. The scores of the distance metric in pixels are to be interpreted in a reverse numerical order, i.e. the lower the score, the higher the accuracy and precision of the participant. A total of 18 participants in three groups were tested on a total of 48 images. Analysis was performed as per the proposed methodology on a total of 864 lines (18 participant lines x 48 images).

4.1.1 Analysis of precision

Out of all 48 images, one way analysis of variance (1-WAY ANOVA) revealed statistically significant differences amongst the three groups ($P<0.05$, $F>F_{crtical=}3.21994$) in 25 out of 48 images. On pairwise comparison, consultants were significantly more precise than trainees (Seniors, Juniors or both) in 14 images. Seniors were significantly more precise than Juniors in nine images. Table 4.1 highlights this data with statistically significant results. Figure 4.2 is a flow chart summarising the precision results of the participant groups in the statistically significant images and figure 4.3 describes the distribution of all the statistically significant pairwise comparisons.

Seniors were significantly more precise than consultants in five images. On direct review of these images it was appreciated that in three cases trainees, albeit more precise, were drawing their lines in a different location on the image. They had the wrong tissue plane. In a further two images, there were significant differences of opinion between consultants as to the location of the ideal tissue plane resulting in orthogonally oriented lines. This seems to generate much larger measures on the modified Hausdorff analysis.

4.1.2 Analysis of accuracy

The differences in accuracy between the senior and junior trainees reached statistical significance in favor of seniors in 14 images whereas junior trainees were statistically more accurate in eight images (Table 4.2 and Figure 4.3).

4.1.3 Duration per participant

Each participant’s time required to complete the exercise is recorded in Table 4.3. The average time taken by Consultants was the greatest amongst the three groups. Average seniors’ time was shorter than consultants’ with juniors’ time being the shortest. These differences reached statistical significance on one-way ANOVA and remained significant on pairwise comparison with a two-sided student’s t- test between consultants and both levels of trainees (C vs J and C vs S). The time taken by senior and junior trainees (J vs S) was however not significantly different (Table 4.3).

4.1.4 Trainee Experience

There was considerable variability in the self-reported experience with laparoscopic and open right hemicolectomies done by trainees (Tables 4.4A and B).

4.1.5 Length of lines

Each participant’s lines drawn on each image were also analysed by the modified Hausdorff distance calculation formula. The average line length of consultants was statistically longer as compared to lines drawn by the trainees (Table 4.5 A and B).

The study took one year to complete (March 2013-Feb 2014). The timeline of the project is elaborated in Table 4.6.

Table 4.1 : Selected images showing comparison of precision between consultant and trainee groups.

Image**	Mean Distance±S.D in pixels			ANOVA(P)*	C vs S (P)	C vs J(P)	J vs S(P)
	Consultant	Senior	Junior				
7	445±221	1778 ±771	1710±707	1.93E-07*	5.55E-06*	4.2E-06*	0.800914
11	597±271	1155±657	1470±694	0.0006*	0.007*	0.0002*	0.212
42	853± 496	2080±1061	1532± 551	0.0002*	0.0006*	0.0006*	0.09
46	853±309	1382±723	1243±611	0.04*	0.01*	0.01*	0.57
22	932±514	1907±731	1770±761	0.0005*	0.0003*	0.001*	0.62
5	969±398	1848±627	1105±335	0.00001*	0.0001*	0.3	0.0007*
23	973±442	1614±927	1722±1089	0.047*	0.02	0.02	0.77
8	1009±409	1560±700	2111±1259	0.0056*	0.016*	0.0055*	0.152952
2	1056±402	2021±1139	1254±616	0.003*	0.006*	0.3	0.03
17	1152±351	2137±807	1484±520	0.0001*	0.0004*	0.05	0.015*
6	1197± 384	1996± 661	2014± 787	0.001*	0.005*	0.001*	0.34
27	1497±669	2114±686	2344±736	0.005*	0.01*	0.01*	0.38
19	1688±830	2500±1201	2732±1081	0.02*	0.04	0.006*	0.58
39	1282±822	2143±669	1178±602	0.0008*	0.004*	0.691	0.0002*

14	1364±574	952±511	865±584	0.04*	0.047	0.026	0.667343
41	989±446	847±441	1390±301	0.002*	0.36	0.01*	0.0007*
9	1072±563	608±210	1973±945	4.707E-06*	0.008*	0.00427*	0.00006*
25	1115±535	829±350	1290±523	0.04*	0.09	0.09	0.009*
47	1205±754	782±228	1450±660	0.01*	0.05	0.05	0.001*
13	1270± 591	770±294	1743±774	0.001*	0.008*	0.121	0.002*
20	1680± 935	996±507	2183± 1158	0.005*	0.01*	0.2	0.001*
18	1684 ±1074	1297± 1455	3833±753	2.88E-07*	0.264	0.00009*	5.522E-06*
40	1783± 733	875±279	1925±748	5.77E-05*	0.002*	0.61	0.0000785*
36	1967±1158	1121±386	1497±674	0.02*	0.01*	0.1	0.07
26	2191±801	1566±412	2489±1222	0.02*	0.01*	0.43	0.01*

*statistically significant difference with Sidak's correction for multiple comparisons ($P < 0.017$)

** Images are arranged in descending order of consultant precision.

SD= Standard Deviation

Table 4.2: Images showing comparison of accuracy between trainee groups.

Image	More accurate	Mean distance in pixels ±SD		P value(Accuracy)
	S or J	S vs C	J vs C	
1	S	1237±770	1564±817	0.08

2	J	1512±903	1171±517	0.03*
3	S	1793±815	2503±911	0.0008*
4	J	1502±711	1359±629	0.47
5	J	1503±729	1106±450	0.0007*
6	S	1552±663	1698±748	0.38
7	S	1218±883	1390±774	0.34
8	S	1400±897	2229±873	0.001*
9	S	866±329	1453±686	0.00003*
10	S	1511±823	1822±852	0.11
11	S	1083±483	1196±542	0.35
12	J	1123±763	956±473	0.27
13	S	1019±503	1467±767	0.004*
14	J	1422±494	1360±581	0.6
15	S	721±300	1128±567	0.0003*
16	S	1428±906	1597±893	0.4
17	J	1712±702	1518±482	0.17
18	S	1699±1218	3148±1258	4.65E-06*

19	S	2250±1001	2260±1226	0.12
20	J	2099±1200	1861±1139	0.39
21	J	1251±545	1159±509	0.46
22	S	1385±668	1540±897	0.42
23	S	1495±915	1509±681	0.94
24	S	1113±552	1453±808	0.04*
25	S	925±424	1414±685	0.003*
26	S	1779±735	2330±903	0.005*
27	J	1923±650	2006±664	0.5
28	J	1745±980	1285±570	0.01*
29	J	2260±1189	2550±1060	0.27
30	S	1969±800	1807±848	0.4
31	S	2117±1287	2482±1158	0.2
32	S	1512±750	2129±1085	0.006*
33	J	2655±803	2179±1121	0.04*
34	S	2870±1509	2389±1387	0.16
35	J	1557±779	1597±816	0.83

36	J	2375±1036	1648±843	0.001*
37	J	1659±704	1697±776	0.82
38	J	1923±930	1491±721	0.03*
39	J	2699±951	1335±783	0.00000000638*
40	S	1552±814	1734±777	7.85E-05*
41	S	925±432	1191±444	0.0007**
42	J	1494±716	1310±590	0.09
43	S	1880±858	2337±838	0.02*
44	J	1317±579	985±474	0.009*
45	J	1055±445	961±468	0.42
46	S	1294±412	1376±765	0.57
47	S	1184±751	1553±1034	0.001*
48	S	1134±598	1165±655	0.99

*statistically significant results ($p<0.05$)

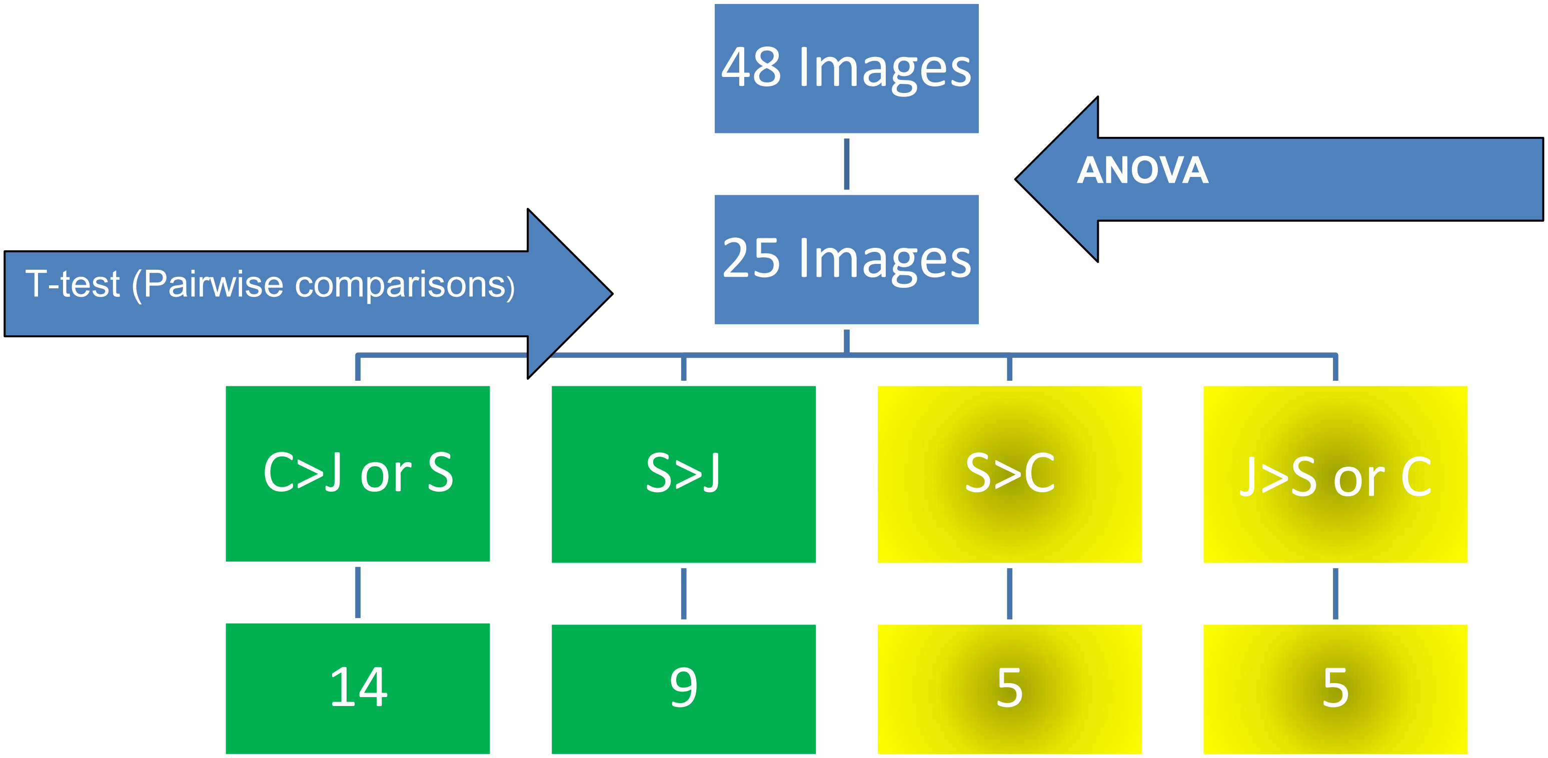


Figure 4.1: Images with statistically significant results based on precision.

The green shaded boxes represent the number of images supporting the hypothesis while the yellow shaded boxes indicate images with conflicting results.

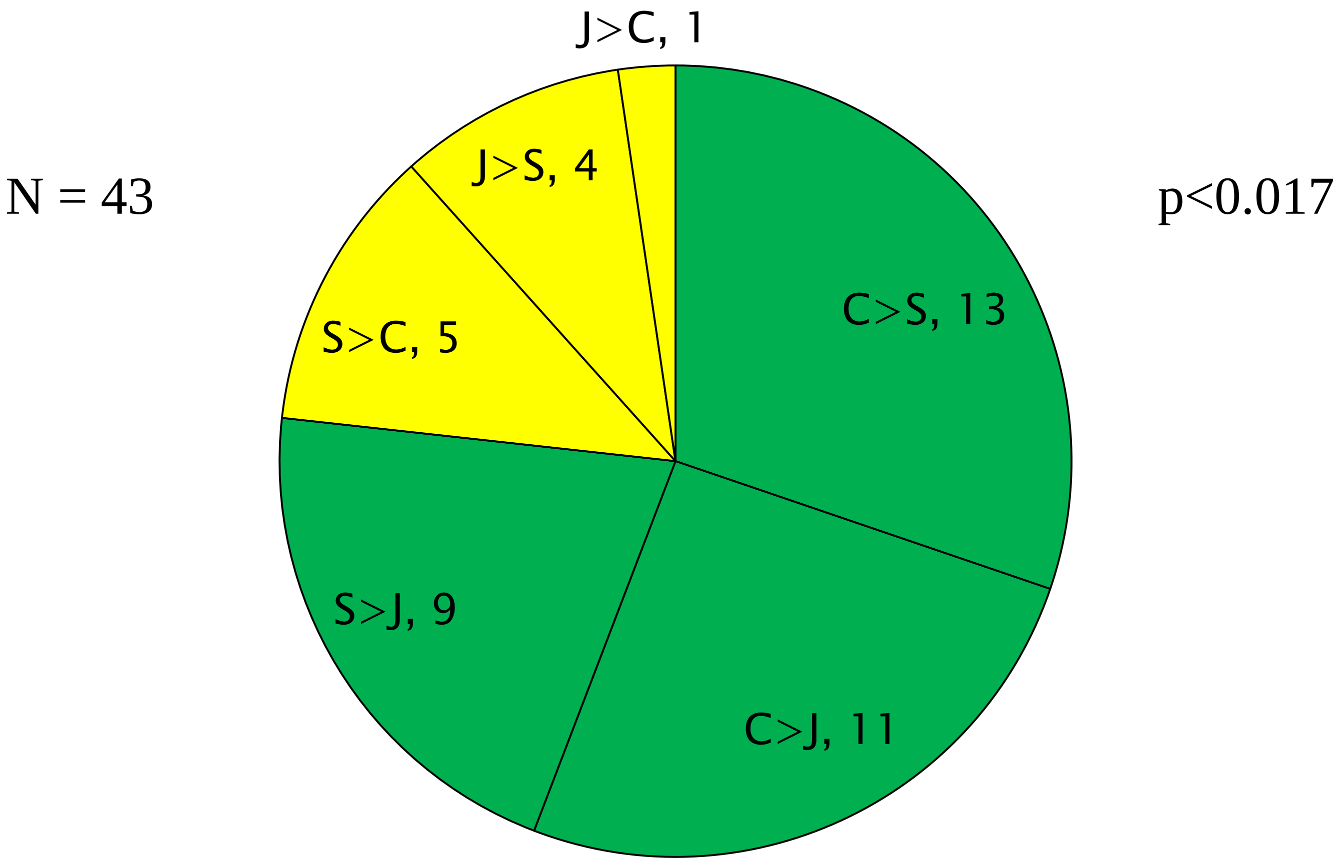


Figure 4.2: Distribution of all the statistically significant pairwise comparisons amongst the three groups.

Three quarters of the results support the hypothesis (indicated by the green area) while one quarter consists of conflicting results (yellow area). When the Sidak's corrected P value is used for multiple comparisons (P<0.017), 43 pairwise comparisons show statistical significance.

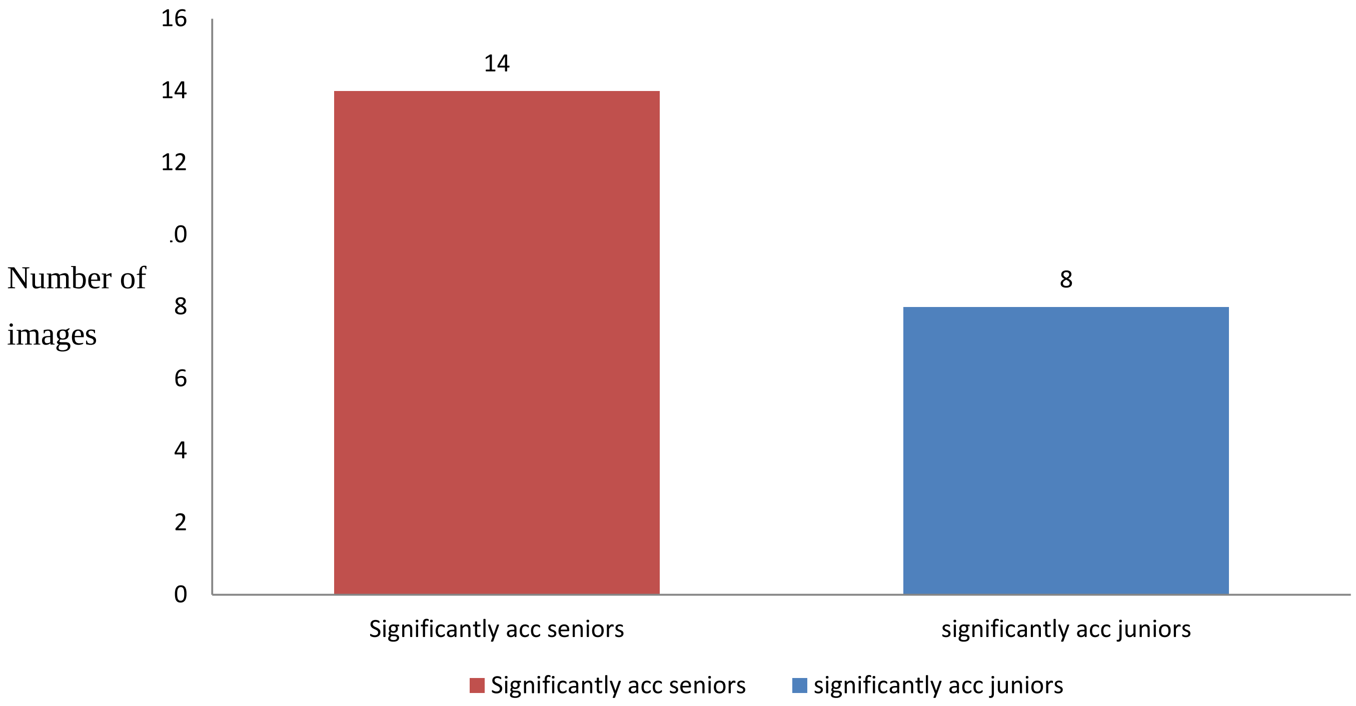


Figure 4.3: Bar chart representing accuracy of trainees.

The y-axis indicates the number of images in which the respective trainee group were more accurate.

Table 4.3: Total time taken by each participant.

Consultants	Time taken	Seniors	Time Taken	Juniors	Time Taken
C1	55	S1	42	J1	29
C2	64	S2	48	J2	32
C3	48	S3	40	J3	31
C4	39	S4	46	J4	28
C5	59	S5	48	J5	36

C6	62	S6	43	J6	35
Mean time (min)±SD	60.75±14.3		40.38±10.7		33±3.54
T-test- C vs S	0.006*	C vs J	0.0009*	J vs S	0.09

Table 4.4 A: Comparing accuracy score of senior trainees to their experience.

Seniors	S1	S2	S3	S4	S5	S6	Mean±SD
Experience	40	30	10	3	2	5	15±16.04
Accuracy	1689.621	1650.319	1313.231	1518.531	1634.828	1658.104	1577±131.3

Table 4.4 B: Comparing accuracy score of junior trainees to their experience.

Junior Trainees	J1	J2	J3	J4	J5	J6	Mean±SD
Experience-number of cases	3	3	15	20	9	10	10±6.69
Accuracy	1703.111	1646	1807	1487	1603	1682	1655±106.91

Table 4.5A: Length of lines drawn by participants for each image.

Image	senior	Consultants	Juniors	Image	Senior	C-line	J-line

	lines				Line Length	Length	Length
1	185.5787	240.9684	251.9261	25	364.5669	500.452	372.0759
2	239.086	356.7852	323.3773	26	332.3228	517.3985	375.1137
3	304.664	523.4079	350.7598	27	341.8356	576.6979	422.4217
4	427.6495	582.3615	466.2588	28	338.4264	515.9522	399.4577
5	431.9403	316.2208	306.0027	29	428.386	460.8107	569.4562
6	361.1794	365.2964	340.6466	30	311.3882	591.1223	370.2004
7	385.5398	331.1948	406.2185	31	328.6989	611.3787	480.4163
8	309.1987	331.0611	409.9014	32	379.6094	393.9628	351.7307
9	368.001	462.5503	403.3229	33	341.3999	584.9601	411.7855
10	432	533	487	34	265.5962	465.0855	394.1451
11	278.3433	442.06	262.6241	35	353.39	432.234	388.5816
12	296.12	363.6632	399.64	36	260.0219	491.7312	394.854
13	450.0211	621.224	463.6962	37	280.9558	456.7309	336.1078
14	409.0713	702.437	428.2759	38	346.1049	428.6113	365.1958
15	410.3743	437.088	512.389	39	379.7473	700.1542	453.3542

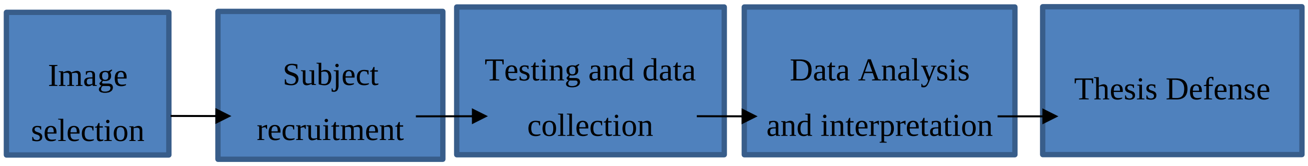
16	355.0476	569.69	327.4889	40	225.0254	421.7729	361.6587
17	320.9317	548.3923	295.9106	41	184.951	209.1826	278.2078
18	244.3948	407.85	324.7139	42	274.115	362.75	368.6424
19	391.5561	491.085	302.9425	43	314.4082	381.2255	346.0547
20	238.431	316.25	262.4664	44	207.1525	298.6695	237.7277
21	467.6933	567.203	391.3248	45	251.2541	371.6456	323.5073
22	439.4454	458.83	281.3172	46	256.2958	484.12	417.723
23	303.4169	434.87	359.7755	47	356.3475	423.9026	438.0473
24	341.2639	485.981	394.5471	48	244.0614	348.783	263.3926

Table 4.5 B: Comparing average line length drawn by consultants, seniors and juniors.

Consultants ± SD	Seniors ±SD	Juniors ±SD
456.6418 ± 72	372.3413 ± 109	328.271 ± 71
Pairwise t-test for line length		
C versus S	C versus J	J versus S
1.53174E-12	1.16679E-07	4.05259E-05

Table 4.6: Timeline of Research Project

May	**June**	**July**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**	**Mar**	**April**
13	**13**	**13**	**13**	**13**	**13**	**13**	**13**	**14**	**14**	**14**	**14**



5 Discussion

5.1 General Overview

This study was a robust attempt to validate the pilot study. Previously done studies have shown that accuracy and precision based analyses are a valid way to assess surgical performance (2, 3, 4, 16). This study increased the number of participants and images. Secondly an expert panel was employed to select the images from a large archive of images. This gave the investigators considerable extent of leverage in selecting the final set of images (48 out of 1000+ images), which again added to the validity of this study.

Though the sample size was still not very large, two more participants per group were recruited as compared to the pilot study and this was more than what was required by the power analysis. It was a reasonably heterogeneous sample, which attempted a resemblance with North American general surgeon population by recruiting surgeons of variable experience and training.

It was ensured that all the participants received the same subset of instructions regarding each image. These instructions were limited to orienting the participant to the procedure and did not have any hints or cues pointing towards the correct tissue plane.

5.2 Significance

There is growing evidence of improved oncologic outcomes associated with adherence to dissection along tissue planes. This has been demonstrated clearly for rectal cancer resections (1). There is compelling evidence to suggest this is also true for colon cancer surgery and hepatobiliary surgery (5, 6). This is a unique exploratory study which to the knowledge of the authors represents a first attempt in designing a simulation based tool to assess a surgeon’s ability to identify tissue planes, which may have important implications with regards to patient outcomes. There has been sufficient evidence in literature pointing towards the importance of visuospatial abilities in surgical task performance and also that with repeated practice, performance on visuospatial tasks can be enhanced (7, 8). These findings can be appreciated in this study as well with consultants performing the best in terms of precision. Literature on simulation based surgical education points towards its importance in enhancing trainee performance in actual clinical settings (9, 10, 11) and this study is a successful attempt towards designing an assessment tool—the first step towards meaningful simulation.

Current training in surgery is going through drastic changes with restrictions in work hours because of patient and trainee safety issues (12). This has resulted in the reduction in the amount of exposure the trainee gets in the operating room. Simulation can play a very useful role in narrowing this gap and is on the top agendas of most North American surgical education and accreditation societies (13).

Current simulation models focus on overall performance and the assessment metrics employed by these simulations are not very content specific (10, 11). In addition, these simulation models have room for improvement in actually simulating the real operative environment. This study is an attempt towards the development of a meaningful simulation tool which attempts to develop a valid assessment metric for a very content specific surgical skill—an ability which most simulation tools lack (10). Specifically pertaining to visual perception, literature suggests the need for gauging performance on very specific visuospatial tasks to assess a surgeon’s visuospatial ability (15). This study is a step in that direction.

This assessment tool once developed and enhanced further has the potential to narrow gaps in trainee exposure and experience and can ultimately be part of the standard general surgical curriculum for various North American surgical training programs.

5.3 Criticism

An interesting finding encountered is the large variability in trainee experience with right hemicolectomy. Some of the junior trainees have more experience with this procedure then the seniors and both groups have large standard deviations. This assessment tool measures a unique surgical ability which is not necessarily acquired specifically by exposure to a right hemicolecotmy. Identifying the surgical tissue plane between the mesocolon and the reteroperitoneum is a skill which is very commonly employed in a large number of abdominal procedures and the skills learned are probably transferable from one procedure to another. Reported experience with right hemicolectomy does not necessarily reflect total operative experience on abdominal procedures. In order to evaluate the impact which trainee experience has on the analysis the senior and junior trainee groups were re-divided based on the trainee's specific experience of a right hemicolectomies. The results of this diminished the statistical significance of the difference in the measures obtained for all three groups, thus potentially adding an element of contamination to the data and did not contribute positively towards the analysis. This supports findings from earlier studies indicating that surgical trainees with similar experience might have very high variability in their learning curve and this has been demonstrated specifically for tasks involving visuospatial abilities (17). It would seem that surgical trainees are using spatial reasoning to identify tissue planes rather than just memory of the specific procedure.

One of the biggest criticisms for this study would be its content validity. Is a surgeon's ability to identify tissue planes being accurately measured? This methodology is based on the assumption that this is an exclusively visuospatial task but it is well known that all surgical tasks and broadly speaking all human tasks fall into overlapping domains of human intelligence, visuospatial and psychomotor skills. It can be seen in actual clinical settings that surgeons tend to use refined manual techniques to expose the tissue such as to improve their visual stimulus and thus this manual dexterity does tend to play a role in facilitating dissection in the correct tissue plane. The pictures used in this study are mostly images in which the tissue plane is already exposed. Therefore it can be confidently stated that this tool does assess the visuospatial ability required to identify a tissue plane. Also exposing tissue is a skill which is more easily acquired through hands on teaching in the operating room where the instructor can actually guide the hand of the trainee to teach them this tactile perceptual skill, unlike visual perception which is often learned by verbal communication and anecdotally. Thus this tool is measuring a unique surgical ability which has not been measured previously and attempts to teach it have also been nonspecific.

Junior trainees’ less accurate performances and their large variability in precision can be not only because of failure to recognize the tissue plane but a general lack of familiarity with the anatomical details of the procedure. This is a potential confounding factor in this study. Junior trainees in this study already had some exposure towards a right hemicolecotmy (at least three operations in the case of the least experienced) and at this level one can presume that a trainee would have a basic understanding of the procedure. However, comparisons between senior trainees and consultants tends to overcome this shortcoming, as senior trainees who have at least three years of intraoperative experience still underperform on precision compared to consultants.

5.4 Unexpected Findings:

5.4.1 Precision

When individual images are considered, the senior trainees are statistically more precise as compared to consultants in five images. These are unexpected results since it was expected that most images would show the consultants being the most precise group. These discrepancies can be explained as follows.

5.4.1.1 Senior Trainees more precise but less accurate

A detailed look at these five images showed that in three of them the seniors had drawn their line consistently at a spot which was clearly away from the actual plane of dissection. Thus they were precise but inaccurate.

5.4.1.2 Disagreement of Consultants

In two of these five images there was a significant disagreement amongst the experts regarding the location of the tissue plane, thus increasing the consultant variability and decreasing precision. This resulted in consultants drawing lines which were orthogonally oriented. This seems to generate much larger measures on the modified Hausdorff analysis. In the future more input from consultants can be obtained while selecting images to overcome this problem.

5.4.2 Accuracy

The accuracy data also tends to point towards junior trainees being more accurate than seniors in a few images and this is the opposite of what was expected. This can be due to several reasons:

5.4.2.1 Difference in innate visuospatial skills

Literature review shows that visuospatial performance in the operating room varies with innate visuospatial skills. According to Keehner *et al.* (7), this ability remains discriminative even at expert level but the extent of the difference does tend to diminish in some ways. Could juniors who scored better than seniors in certain images be visuospatially more gifted? Performance of participants on standard battery of visuospatial tests was not recorded before they went through the assessment as many earlier studies had done (7, 8). Otherwise it would have been interesting to note if visuospatial abilities correlate strongly to performance specifically at the trainee level as had been suggested by these studies. Also as mentioned above the variability in trainee learning curve could cause some of the junior trainees to be more accurate (17).

5.4.2.2 No gold standard for accuracy

Unexpected differences in accuracy between the two trainee groups can also be attributed to a lack of a clear gold standard for determining accuracy. Consultants are being used to assess accuracy based on their clinical experience but all consultants do not have an absolute agreement on the precise location of the dissection plane.

5.5 Other Findings

5.5.1 Line-length discrepancy

During the study, trainees made their lines much shorter than the consultants. This was attributed to the fact that consultants were surer about the location of the tissue plane in the whole image and thus were drawing bigger lines. The trainees and specifically the senior trainees spend a significant amount of their time in the operating room following instructions from their instructors who ask them to perform specific tasks in a procedure. Thus they intuitively develop a habit of making small specific movements in the operating room. The sub analysis showed that the line length was the greatest amongst the consultants and was also directly proportional to experience, as there was a statistically significant discrepancy in line length amongst the three groups. This line length discrepancy might result in decreasing the variability and accuracy estimate of the modified Hausdorff distance metric. When line length is smaller, the nine equidistant points tend to be tightly clustered together and lines having tightly clustered Cartesian points might give only small differences in the Hausdorff distance. No mathematical way to adjust for this has been developed yet.

5.6 Future Directions

5.6.1 Line Length

As pointed out earlier, differences in length of lines between participants can potentially have confounding effects on trainees' and consultants' precision. In the future, the plan is to give a specific area of the image to participants to draw lines where the expert panel agrees on the existence of the tissue plane. This technique can add uniformity to the length of lines amongst different group members. Another way to counter this would be to add the length of lines into the modified Hausdorff's distance metric such that if a participant draws a line of a greater length, it would decrease that participant's precision.

5.6.2 Most discriminative images

Images with the strongest discriminative power will be used to conduct further studies. These are images where consultant's precision is statistically higher than the other two groups, and the seniors are statistically more accurate than juniors.

5.6.3 Addition of dynamic visuals

A field of interest is the effect which this tool would have if video context is added to images before showing them to participants. The video would include a few seconds into the actual operation before the image is captured. Although the evidence for static versus dynamic images for enhancing a simulation is equivocal (18), this is a field which requires exploration specially if the plan is to simulate actual surgical environment.

5.6.4 Teaching tool

The ultimate goal of this project is to develop a simulation tool for surgical trainees which can help them identify correct tissue planes. This project achieves the first step towards laying down the foundations of a unique and specific assessment tool.

5.7 Conclusions

From results of this study, there is enough evidence to support the hypothesis that this simulation based assessment tool accurately distinguishes surgeon's ability to identify surgical tissue planes, with the more experienced group performing better at this simulation. These results reach statistical significance both in terms of precision and accuracy. The unexpected results in some images could be a result of several factors which include images lacking discriminative power, the consistent inaccuracy of the trainees, the lack of a gold standard for comparing accuracy, line length discrepancy, and innate differences in visuospatial abilities. In the future, the authors will try to find ways to minimize these effects. The ultimate goal of this study is to be transformed from benchmark research into a clinically applicable teaching tool to enhance trainees' learning experience and play a role towards patient safety.

5.8 References

1. Joint Committee on Surgical Training: Guidelines for the award of a CCT in General Surgery (Internet). JCST, London UK. (Date cited : 02/04/14) Available from: http://www.jcst.org/quality_assurance/Docs/cct_guidelines_gs

2. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. Stud Health Technol. Inform. 1998; 50:124-130.

3. Darzi A, Mackay S. Assessment of surgical competence. Qual. Health Care. 2001;10:(Suppl 2):ii64–ii69.

4. Risucci D, Cohen JA, Garbus JE, Goldstein M, Cohen MG. The effects of practice and instruction on speed and accuracy during resident acquisition of simulated laparoscopic skills. Curr. Surg. 2001; 58:230-235.

5. West NP, Morris EJA, Rotina O, Cairns A, Finan PJ and Quirke P. Pathology grading of colon cancer surgical resection and its association with survival: a retrospective observational study. Lancet Oncol. 2008; 9: 857–65.

6. Imamura H, Seyama Y, Kokudo N, Maema A, Sugawara Y, Sano K, *et al*. One thousand fifty-six hepatectomies without mortality in 8 years. Arch Surg. 2003; 138(11): 1198-1206.

7. M. Keehner, Y. Lippa, D.R. Montello, F Tendick, M. Hegarty. Learning a Spatial Skill for Surgery: How the Contributions of Abilities Change With Practice. Appl. Cognit. Psychol. 2006; 20: 487–503. DOI: 10.1002/acp.1198.

8. Kyle R Wanzel, Stanley J Hamstra, Dimitri J Anastakis, Edward D Matsumoto, Michael D Cusimano. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet. 2002; 359: 230–231

9. Dawe SR, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, Maddern GJ.A systematic review of surgical skills transfer after simulation-based training: laparoscopic cholecystectomy and endoscopy. Ann Surg. 2014 Feb; 259(2):236-48. doi: 10.1097/SLA.0000000000000245.

10. Leanne M. Sutherland, Philippa F. Middleton, and Guy J. Maddern. Surgical Simulation A Systematic Review. Ann Surg. Mar 2006; 243(3): 291–300.doi: 10.1097/01.sla.0000200839.93965.26

11. David A. Cook, Rose Hatala, Ryan Brydges, Benjamin Zendejas, Jason H. Szostek, Amy T. Wang, Patricia J. Erwin, Stanley J. Hamstra. Technology-Enhanced Simulation for Health Professions Education: A Systematic Review and Meta-analysis. JAMA. 2011; 306(9):978-988. doi:10.1001/jama.2011.1234.

12. American Academy of Graduate Medical Education: Recommendation of Committee on work hour restrictions (Internet Article). AAGME, Chicago, IL (Date cited 02/04/2012) Available from https://www.acgme.org/acgmeweb/tabid/271/GraduateMedicalEducation/DutyHours.asp

13. Position of the American College of Surgeons on Restrictions of work hours.(Internet).ACS Chicago IL.(Date cited: 02/04/2014) Available from http://www.facs.org/education/statement.pdf

14. Clare Faurie, Mohammed Khadra. Technical competence in surgeons. ANZ Journal of Surgery. 2012 Oct; 82(10): 682-90.

15. Stan Hamstra, Visual perception in surgery: How can its challenges be addressed to improve training and performance? (Internet) American College of Surgeons educational portal for surgical residency program directors. ACS, Chicago IL. (Date cited 02/04/2014) Available from http://www.facs.org/education/rap/hamstra1208.html

16. Martin J, Regehr G, Reznick RK, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br. J. Surg.1997; 84:273–278.

17. Jan-Maarten Luursema , Willem B. Verwey, Remke Burie. Visuospatial ability factors and performance variables in laparoscopic simulator training. Learning and Individual Differences. 2012; 22 : 632–638

18. Keehner M, Khooshabeh P, Hegarty M. Individual differences among users:Implications for the design of 3D medical visualizations. User Centered Design for Medical Visualization. Hershey, PA: Information Science Reference, an imprint of IGI Global; 2008. 1-24 p.

Abbreviated Curriculum Vitae

Name:	Syed Ali
Post-secondary Education and Degrees.	University of Karachi, Pakistan Bachelor of Medicine and Surgery (MBBS). 1999-2003
Honors and	Best resident teacher for surgery clerkship, University of

Awards

Toronto General Surgery Program
2008

Best senior resident teacher for junior residents
2009-2010

CAGS- Best Basic Science Award- A Novel Model for Measuring Surgeons' Visual Perception of Tissue planes
2013

Related Work Experience

Faculty Instructor at the Shulich School of Medicine and Dentistry
Western University
2012-2014

Clinical and Research Fellow, Canadian Surgical Technologies and Advanced Robotics, Western University
2012-2014

Publications:

Talat Chughtai, **Syed Ali**, Phillip Sharkey, Marcelo Lins, and Sandro Rizoli. Update on managing diaphragmatic rupture in blunt trauma: a review of 208 consecutive cases. Can J Surg. 2009 June; 52(3): 177– 181.