Flying Aptitude Tests for Surgeons

Hyunmi Park MBChB MRCS

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

2017

DEDICATION

To God.

To my parents AhnSik Park and YoonSoon Kim, for their abounding and selfless love and support.

To my brother Joonjae and sister-in-law Yujun, for giving me such joy through my loving nephew Ethan and niece Elena.

To my best friend Jihye, as a PhD, she totally understands.

부모님께 이 논문을 바칩니다.

내게 생명을 주시고, 끝없는 사랑과 헌신으로 저를 키워주시고 격려해주신 엄마 아빠. 진심으로 사랑하고 감사합니다.

ACKNOWLEDGEMENT

I would like to sincerely thank my supervisor and surgical mentor Mr Charles Maxwell-Armstrong, who has tirelessly encouraged me and without whom this work would not have been possible.

I am also very thankful to my supervisors Prof Reg Dennick and Prof Eamonn Ferguson for their kind support.

I would like to thank the staff at the Officers and Aircrew Selection Centre RAF Cranwell, especially Air Commodore Bill Coker and Psychologist Dr Margaret Bailey.

I would like to thank all the staff at the Trent Simulation and Clinical Skills Centre.

Finally, I thank all my students who have shaped me into being a better teacher.

TABLE OF CONTENTS

ABSTRACT

Aims: To test whether the Flying Aptitude Test, the Royal Air Force's selection aptitude test, can be used for surgical skills testing by correlating it to both open and minimally invasive surgical simulated tests. The introduction of such test at the onset of post graduate training could offer guidance and encouragement on a career in surgery.

Methods: The Flying Aptitude Test used for the selection of British Military Pilots was undertaken by 243 medical participants. The aptitude domains tested included: Psychomotor, Verbal, Attentional, Spatial aptitudes & Short-Term Memory. Results were correlated with performance of open Basic Surgical Skills (BSS) and Laparoscopic Simulator (Lap Sim) skills in simulated environments. Medical students (n=211) encompassed 86.6% of those recruited and the remainder were doctors in training. Correlation analyses were carried out on the undergraduate participants only to maintain a completely uniform group of novices from both surgical and military aptitude experience. Data on demographics, use of computer games, self-rating scores and feedback form were also analysed.

Results: n= 243 (52.3% female). Mean age 24 years (range18-39). 230 participants undertook the computer based Flying Aptitude Test of which 199 were medical students with a mean score of 51.64% (16-96% SD=14.27). Total mean Lap Sim time was 737 seconds (259-2290sec SD=313). Twenty-six participated in the BSS with a mean score of 74% (16-97%, SD=23). There was statistically significant correlation between the Flying Aptitude Test and the Lap Sim data (undergraduates n=153 Pearson r=-0.275; p<0.001) with the highest correlation in the Psychomotor domain (r=-0.300; p<0.001). There was even greater correlation between the Flying Aptitude Test and BSS tests (undergraduates n=20 Spearman's r=0.464, p=0.04) with the Spatial Reasoning aptitude having the highest correlation (r=0.540, p=0.014). Lap Sim & Flying Aptitude Test data correlation was greater in females but for the BSS data, the correlation was greater in male, but this difference between the genders was not statistically significant. Positive correlation was seen in the use of computer games and the Flying Aptitude Test, which was higher in males. There was a marked difference in the self-rating results between the genders, with female participants reporting an unfounded lower expectation of their own performance.

Conclusions: This study shows a statistically significant correlation between the validated Flying Aptitude Test scores in both open and laparoscopic simulation tests. This study has shown an equally good performance from female medical students compared to their male peers in the Flying Aptitude Test as well as the Laparoscopic and Basic Surgical Skills Tests in this study. A surgical aptitude test such as the Flying Aptitude Test could be potentially incorporated into early post graduate training to inspire graduates into a career in surgery. Such aptitude test may encourage self-actualisation and empowerment of female trainees into believing in their own potential technical ability and challenge the gender gap in the speciality.

Chapter 1: INTRODUCTION

'Born to be good, train to be great'

JC Hall, BJS 2003[1]

1.1 Background

My personal experience as a UK graduate in surgical training and as a teaching fellow for undergraduate medical students has influenced my interest in medical education both in the undergraduate and post graduate setting. The idea for this study stemmed from our previously published work[2] comparing British military Harrier pilots and undergraduate medical students on four previously validated[3] simulated laparoscopic surgical tests. Both groups were surgical novices and did not have any prior experience in laparoscopic surgical instruments. The first group of participants was made of UK Harrier pilots who had been selected into training through a rigorous three-day selection process which includes the computer based RAF Flying Aptitude Test. The Flying Aptitude Test includes the testing of aptitudes such as psychomotor dexterity and hand-eye coordination, which are aptitudes known to be of importance and relevance in surgery, especially minimally invasive surgery. Each Harrier pilot in the study also had more than 500 hours of flight simulator experience.

The second group consisted of self-selected undergraduate Nottingham University medical students and the third group comprised of non-medical undergraduate students.

There was a statistical significant difference in the speed between the groups. The Harrier pilots were more than three times as fast and made zero errors. The Harrier pilots' Appendicectomy task was completed in 28.7 seconds and the medical students' mean time was 95.5 seconds. The speed had not compromised accuracy with the Harrier pilots excelling in their first time at the laparoscopic simulator. The study raised questions on what the potential different qualities within both set of surgical novices and what different methods of selection and training were used that produced such different results.

Professor of Orthopaedic Surgery Charles SB Galasko[4] responded to a BMJ letter[5] comparing the aptitude of pilots and surgeons. He stated that the competencies required to be a consultant surgeon in the modern NHS are many, and include; learning skills, personal effectiveness, communication skills, teaching skills including appraisal and assessment, in addition to the required knowledge of basic sciences, clinical sciences, clinical skills, surgical skills and manual dexterity and decision making and judgement. He emphasised that the development and competence in all of the above skills are very important[6], but good technical skills being the bedrock of successful surgery[7].

The *Good Medical Practice* report published by General Medical Council[8] states that the standard of care and respect for human life should be justified by the following four domains;

- Knowledge, skills and performance
- Safety and quality
- Communication, partnership and teamwork
- Maintaining trust: being honest and fair to patients and colleagues

This study does not cover all of the many duties of a good surgeon, outlined above, but investigates the first skill listed, which is that of providing a good standard of surgical practice to our patients. Surgeons should demonstrate and maintain competence in their area of clinical practice by keeping up their skills and knowledge which, where relevant, makes appropriate use of simulation to support learning of new procedures.

This introductory chapter will provide essential background information on the Flying Aptitude Test used by the British Armed Forces (mainly the Royal Air Force) to select their candidates, on surgical training and surgical simulation tests. The chapter ends with the Aims and Objectives of this thesis.

1.2 Flying Aptitude Test

1.2.1 History of the Flying Aptitude Test

The evolution of the aptitude testing in the RAF started at the beginning of the Second World War. The selection of aircrew into the RAF prior to the 1940s consisted merely on an unstructured interview, where the serving officers had no further brief than to find the 'right types' as the selection panel[9].

Bailey M[9], the head of psychology at the Officers and Aircrew Selection Centre (OASC) RAF Cranwell published that this method of selection, without the guidance on the relative merits of personality, attainment and skills and without the technical aides to measure them resulted in the very high incidence of pilot training failure. This failure resulted in an unacceptable number of flying accidents, which was very costly and unproductive. The introduction of a systematic testing method into the aircrew selection began when the failure rate of trainees of up to 50% proved to be a major issue at the start of WWII. At the request of the Air Ministry, the first battery of tests was developed in 1941 at Cambridge University. Prof FC Bartlett founding the Training Research Branch recognised the importance of objective selection testing in reducing training wastage rates. By 1944 with the introduction of electro-mechanical coordination tests into the standard battery of tests resulted in the reduction of pilot cadet training wastage from 48% to 25%. The RAF also identified the need to train and select adequate panel members who could cope with test administration and marking. This standardisation was implemented even at selection panel level.

The US Army Air Forces under the direction of Melton[10] during the Second World War also made enormous efforts towards the development and validation of psychomotor tests. Dozens of apparatus were developed and evaluated for psychomotor abilities and administered to more than 600,000 men through their programmes. They used the tests for selection and classification of aircrew personnel as early as 1942. Future publications of reviews by Salvendy & Seymour[11] in 1973 and Adams[12] in 1987 demonstrated how Melton's psychomotor tests were clearly demonstrated to be valid predictors of complex-task performance.

From 1985 a computerised version of the tests was introduced at the OASC (Figure 1) with the help of cheap micro-computer technology.

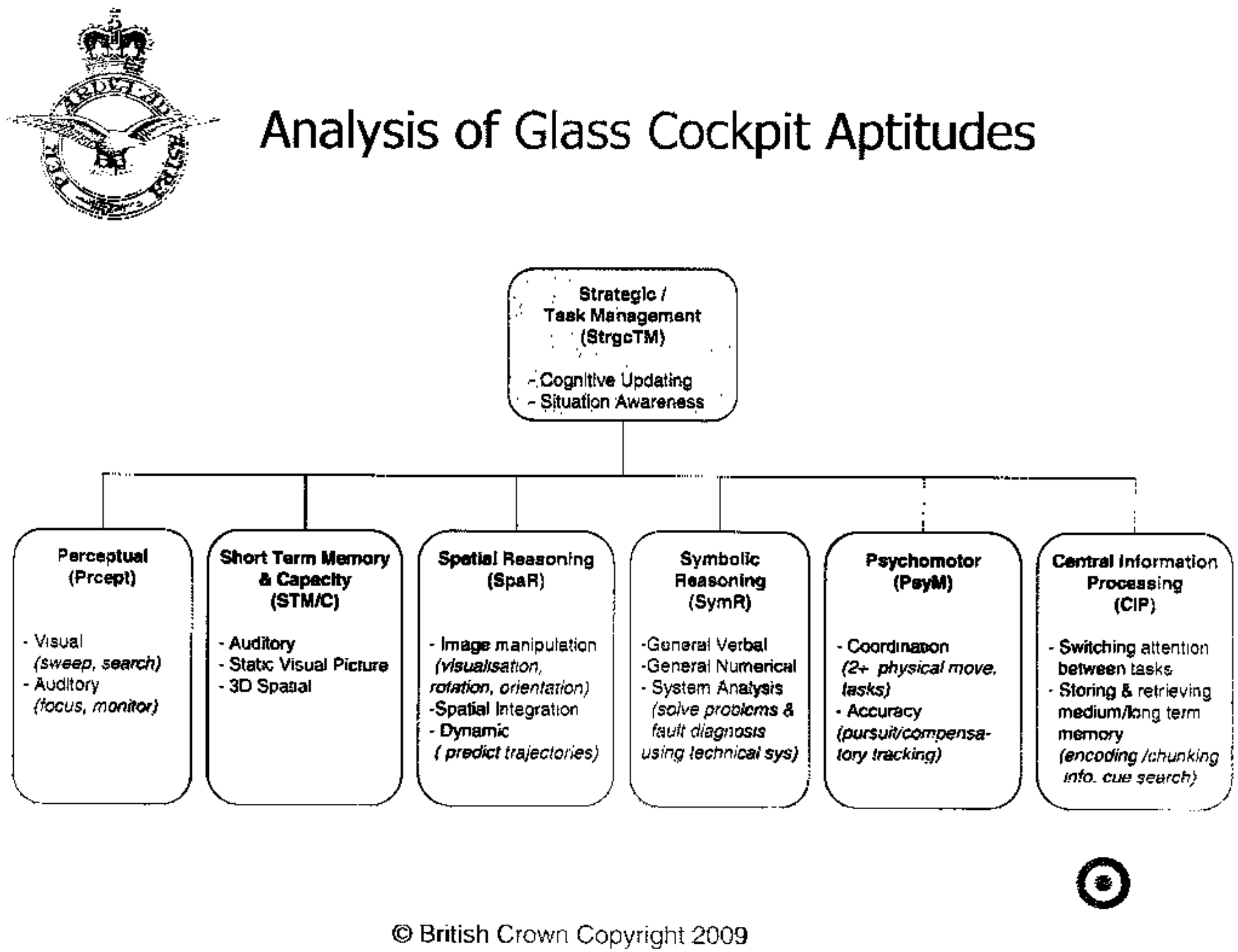
Up to that point no critical changes were made from the original system established during the Second World War. The result was the improved reliability of the tests scores, especially in the test-retest reliability of the co-ordination tests. Computerisation was also welcomed by psychologists who had to collate both the manual records of test results and training data to evaluate the psychometric properties of tests and their predictive effectiveness, which was very time consuming and prone to errors. The evolution of the tests 'The Air Traffic and Fighter Controller Test Battery' was found to have a good predictive power at 0.52 according to the beta values from the regression analysis. In the 1990's the selection process in the RAF shifted from the tests themselves to measuring aptitudes, which meant that the type of tests could evolve with the introduction of new technology but the underlying aptitude being tested would remain.

Figure 1: The evolution of the computers used for the Aptitude Testing by the OASC RAF Cranwell

Sensory Motor Apparatus	Electro-mechanical coordination test
	
1986-1991 Torch Computer	Archimedes Computer 1991-1998
	
PISC PC Computer 1998-2005	2005-present day
	

The aptitudes for each job were identified after a job analysis was performed involving subject matter experts (individuals with a thorough knowledge of the job) evaluating the importance of each domain and the suitability of available tests. Each role (e.g. Air Traffic Controller and Fighter Controllers) was progressively broken down to individual job tasks at operational level. The importance of each job task was then weighted and the aptitudes required for each job task identified. Figure 2 describes the aptitudes within the selection of aircrew at the OASC used in 2009 identifying six distinct areas for testing after an analysis of the Glass Cockpit Aptitudes for pilots.

Figure 2: RAF Glass Cockpit Aptitudes 2009



1.2.2 The Current RAF Flying Aptitude Testing

There are eligibility checks before being able to apply for selection into the RAF as a pilot. These include being a citizen of the United Kingdom, aged between 17.5 and 25 years and meeting certain educational, fitness and health requirements. As for entry into the medical profession, a background check is also necessary [13]. The RAF guidance notes that there is no limit on the number of times an aptitude test may be taken, but that a period of at least one year must elapse between each attempt to minimise the Practice effects, where the candidates may perform differently in subsequent attempts because of familiarity with the tests.

The selection process starts with a series of computer based aptitude tests at the Officers and Aircrew Selection Centre in RAF Cranwell, ensuring that the RAF matches the right candidate to the right role. The aptitude test is followed by a selection interview, occupational health assessment, pre-joining fitness test and further specialist interviews.

Within the Aircrew Branch in the RAF there are other positions such as the Weapons Systems Operator and Remotely Piloted Aircraft System (RPAS) who also are selected using the Flying Aptitude Test. The selection tests for the RAF Pilot Branch are the same as those used for Royal Navy Fleet Air Arm and Army Air Corps flying selection.

The Flying Aptitude Test consists of multiple-choice tests, each of which assesses the candidates' aptitude in a different area and have been divided into the following six aptitude domains:

1. Verbal Reasoning
2. Numerical Reasoning
3. Short Term Memory
4. Spatial Reasoning
5. Psychomotor
6. Attentional Capability

The images of the tests have been provided by the RAF in their "Notes for Guidance of Candidates Attending the Royal Air Force Officers and Aircrew Selection Centre Computer Based Aptitude Testing" document [14]

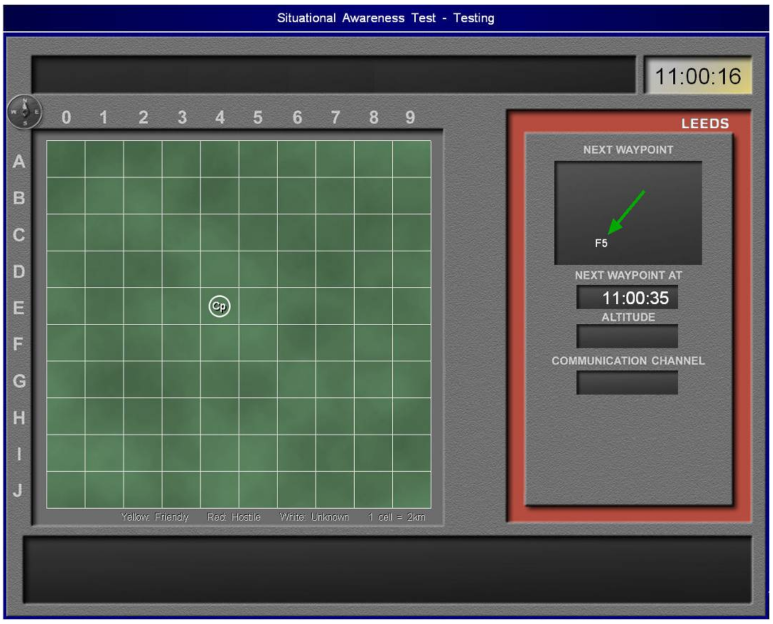
1. **Verbal Reasoning**: Domain referring to the ability to interpret and reason with verbal information which can also be in written format. It is the ability to identify patterns in presented information and to solve problems by combining sensible rules of thumb with a logical approach.

The Verbal Logic Test (Figure 3) assesses the ability to use and interpret written verbal information from different sources to solve problems. The candidate is required to collate relevant information from a number of folders available from a menu system and interpret and analyse verbal information presented in paragraphs. Finally making deductions, inferences and evaluations based on available information.

Figure 3: Verbal Logic Test

The Situational Awareness Test (Figure 4) is a multiple tasks test designed to assess the candidate's ability to collate verbal, numerical, pictorial information to build, maintain and update a mental picture of a complex changing situation to solve problems. The candidate is required to monitor verbal, numerical, pictorial and coded information which can be presented aurally via a set of headphones or visually on screen. The participant has to monitor the changing situation, be aware of past, on-going and possible future activities/events that concern the position and movement of a variety of objects and at the same time answer queries on the changing situation and think about what actions are appropriate to take to solve various problems.

Figure 4: Situational Awareness Test



2. Numerical Reasoning: The ability to use and interpret numerical information using basic fractions, decimals and formulae and understanding and interpreting graphs and tables. There are various tests under this domain but there were not included in this study. Examples of these tests are:

The Airborne Numerical Tests assess the ability to estimate answers to numerical calculations whilst under a degree of time pressure, as demanded in an airborne environment.

The Angles, Bearings and Degrees Test, is designed to assess candidates' judgement of angles and bearings by estimating the angle between two lines and estimating the bearing of an object from a reference point.

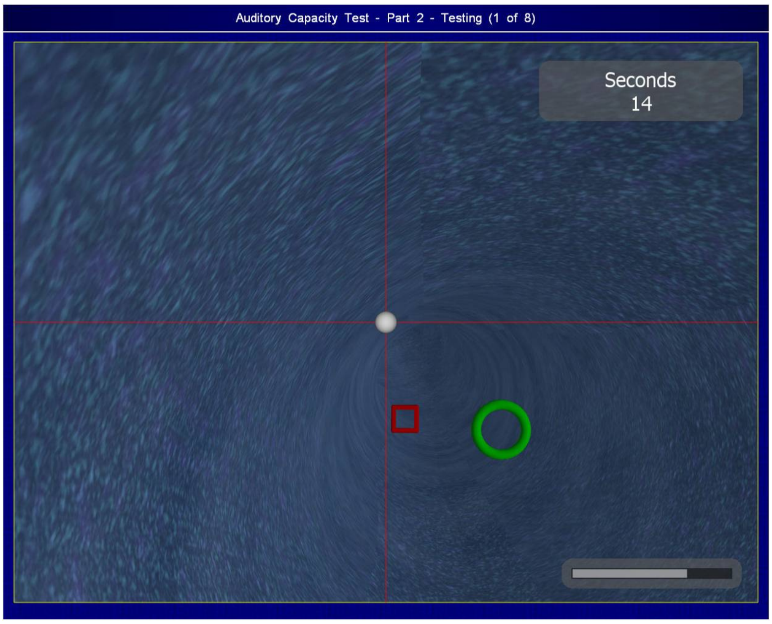
The Mathematics Reasoning Test is designed to assess the ability to solve numerical problems by the interpretation of written descriptions to solve numerical problems using time/speed/distance calculations.

The Numerical Operations Test is designed to assess the ability to work out mental arithmetic problems and mental calculation using addition, subtraction, multiplication and division whilst working quickly and accurately.

3. **Short Term Memory**: The retention of information which then has to be recalled within a short period of time.

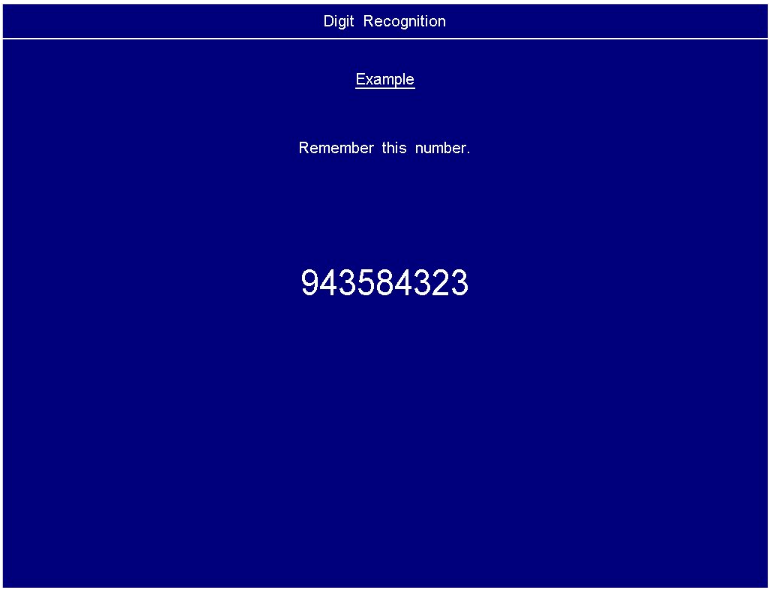
The Auditory Capacity Test (Figure 5) is designed to assess the memory capacity under multiple tasks and in timed condition. The test requires the candidate to deal with more than one task at a time, e.g. change colours and numbers using the keypad provided, control the movement of a ball presented on screen using a joystick and remember a string of figures whilst following aural instructions given over headphones to remember and recall appropriate actions to take regarding the different tasks.

Figure 5: Auditory Capacity Test



The Digital Recognition Test (Figure 6) is a short-term memory test designed to assess the candidate's short-term visual memory by remembering strings of digits of varying lengths and answer questions about the string of digits shown.

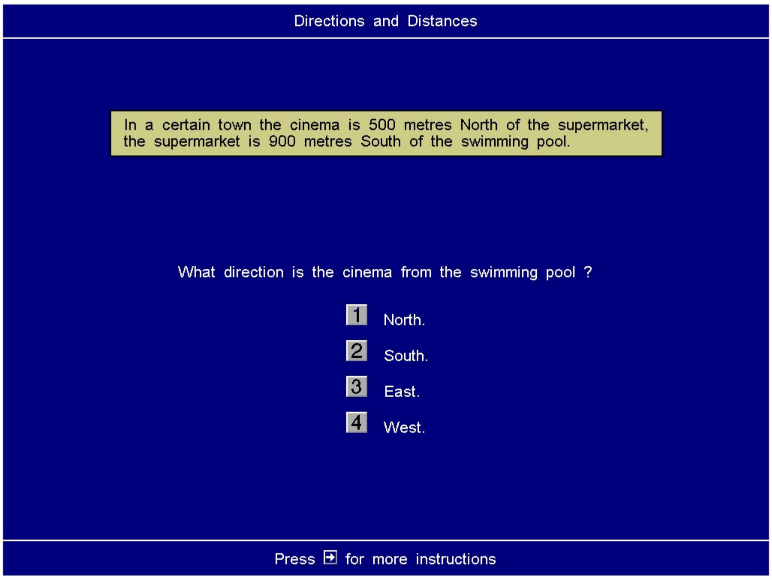
Figure 6: Digital Recognition Test



4. Spatial Reasoning: The ability to form mental pictures and manipulate spatial information in the mind. The ability to form three-dimensional representations (mental pictures), and manipulate diagrammatic/pictorial information is sometimes called 'the mind's eye'. Image manipulation requires visualisation, rotation and orientation of the presented information.

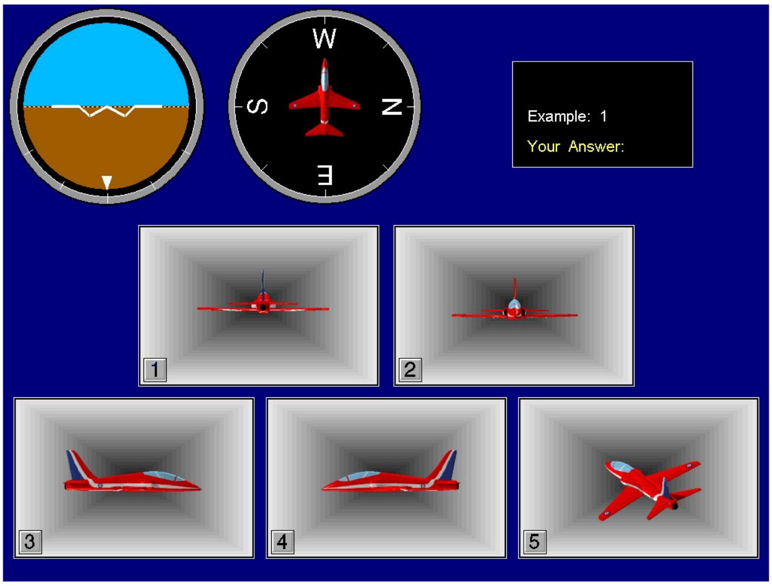
Interpretation of Three-dimensional spatial image descriptions and its spatial relationships is tested in the Directions and Distance Test (Figure 7). The candidate is required to interpret written descriptions of the distance and direction between objects and answer queries on the spatial relationship between different objects.

Figure 7: Directions and Distance Test



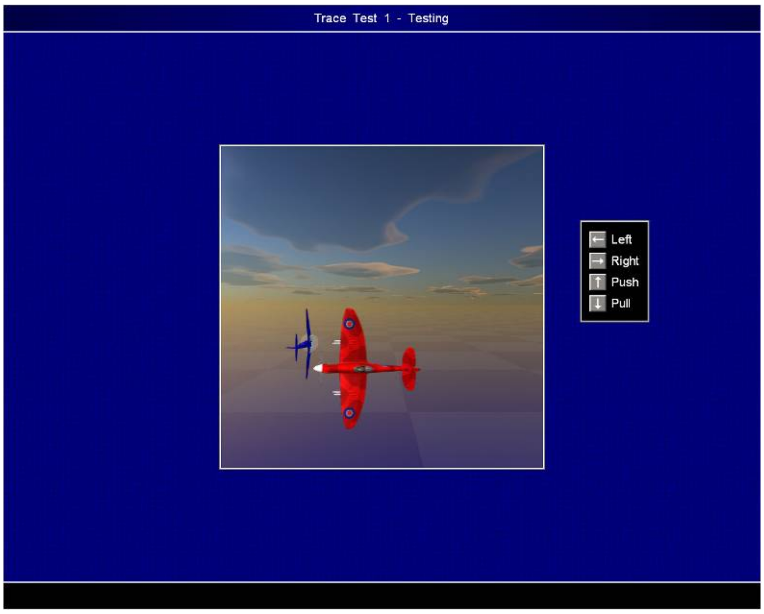
The Instrument Comprehension Test (Figure 8) is a spatial test designed to assess the participant's ability to visualise using pictorial, numerical and verbal information by inspecting instrument readings to visualise the orientation of an aircraft.

Figure 8: Instrument Comprehension Test



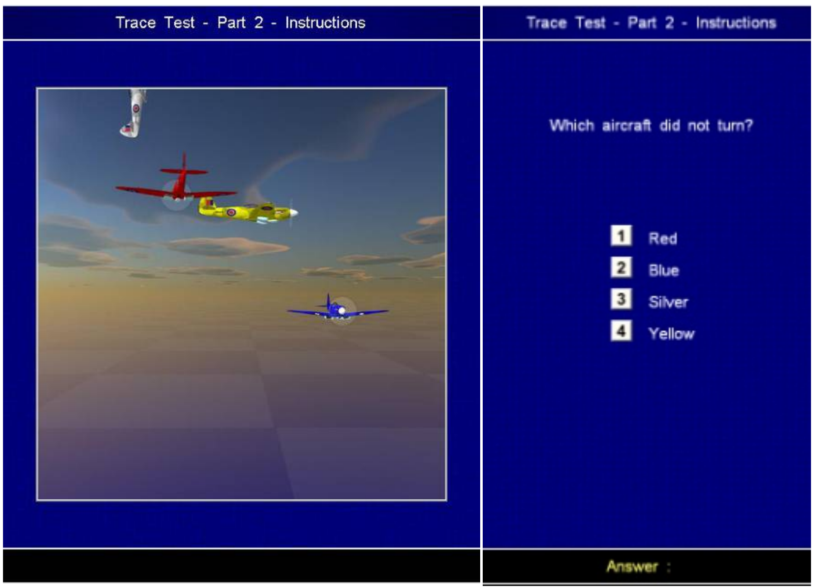
The first Trace Test (Figure 9) is a spatial test designed to assess the ability to orientate in three-dimensional space. The participant is required to perceive the changing orientation of a moving aircraft from a different perspective/direction and then interpret pictorial information and identify the change in orientation of a given aircraft.

Figure 9: Trace Test 1



The second Trace Test (Figure 10) is designed to assess the ability to remember the movement of objects in three-dimensional space. The participant is required to watch short dynamic scenarios involving the movement of a number of aircraft and then recall the movement and location of moving aircraft in three-dimensional space.

Figure 10:Trace Test 2

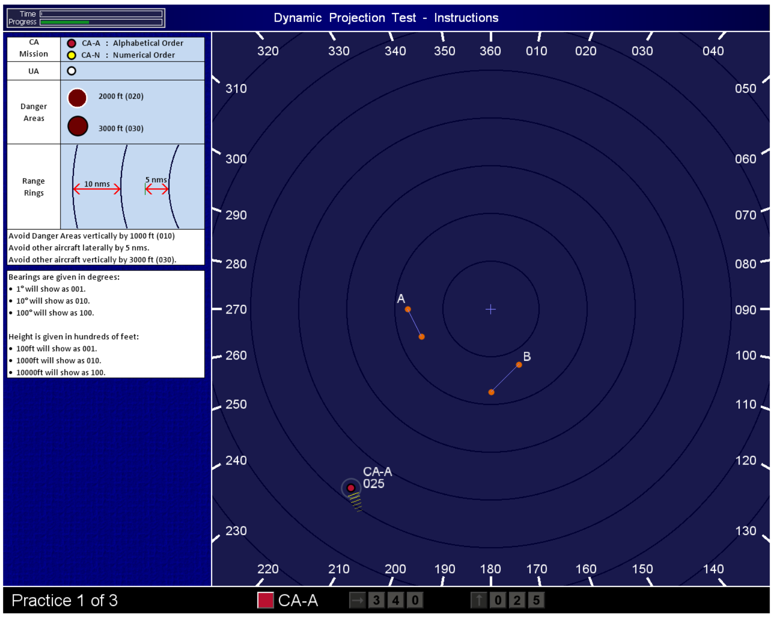


The Spatial Integration Test (Figure 11) is designed to assess the ability to collate information provided by a number of two-dimensional displays to form a 3D air/ground picture. The candidate is required to interpret and integrate information presented from various 2D aerial viewpoints about the location and surrounding objects as well as the location of moving aircrafts. The candidate also has to form a 3D air/ground picture of the location and surrounding objects/aircraft, as seen from different directions/angles.

Figure 11: Spatial Integration Test

The Dynamic Projection test (Figure 12) is designed to assess the ability to interpret and direct the movement of objects within a 3D, dynamic environment. The candidate is required to direct moving objects towards specific points and other objects using bearings whilst avoiding moving objects and obstacles.

Figure 12: Dynamic Projection test



5. Psychomotor Coordination: The ability to demonstrate physical coordination, commonly hand-eye or hand-eye-foot, incorporating two or more physical moves.

The Rapid Tracking Test (

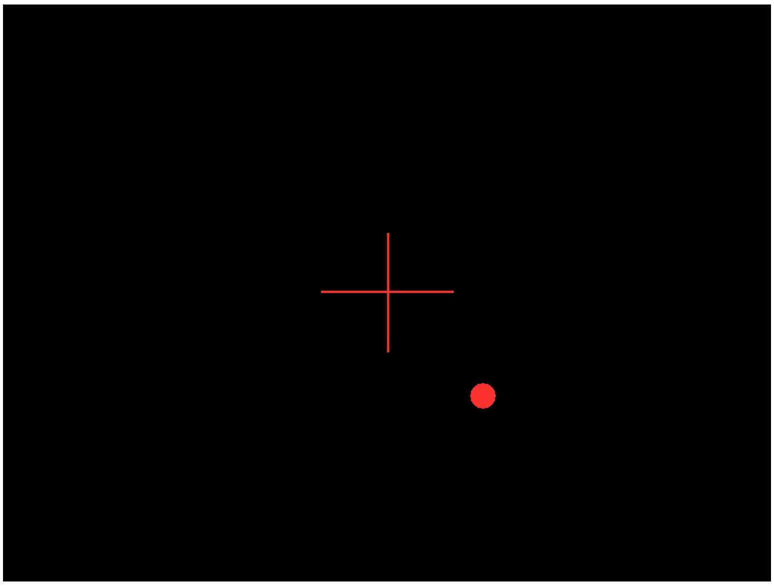
Figure 13) is an eye-hand coordination test is designed to assess the ability to track and target objects. The candidate is required to use a joystick to track moving and stationary objects from a location that is moving continuously and also predict the movement of obscured objects.

Accuracy of the psychomotor task measured in pursuit or compensatory tracking tests. The Sensory Motor Apparatus Test (

Figure 14) requires eye-hand-foot coordination where the candidate is required to use a joystick and foot pedal to control the vertical and horizontal motions of a moving circle (i.e. red dot) and keep it as close as possible to the centre of the cross-hair. This test has not changed since its introduction more than half a century ago but is in computerised since the 1980s (see first picture of Figure 1).

Figure 13: Rapid Tracking Test

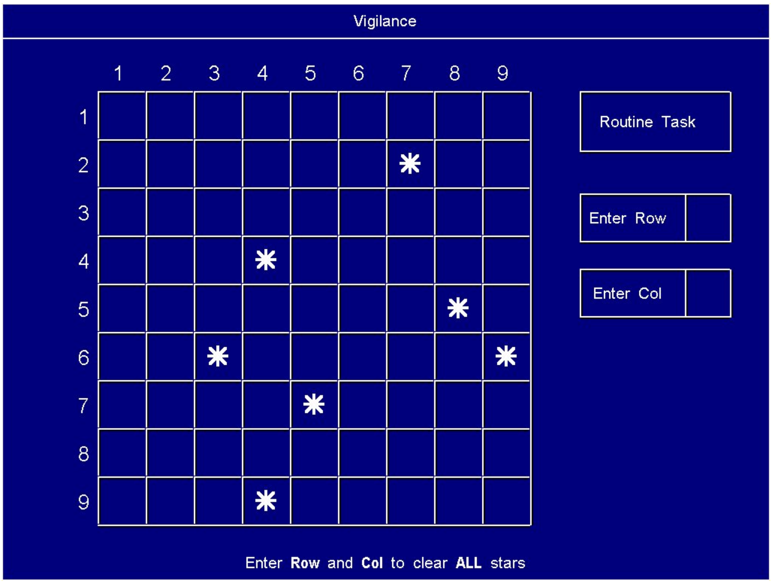
Figure 14: Sensory Motor Apparatus Test



6. **Attentional Capability** also described as **Work Rate**: The ability to switch attention between tasks (commonly known as multi-tasking). A broad range of abilities are tested including the ability to deal with multiple tasks involving auditory and / or visual information as well as storing and retrieving medium/long term memory; concentrating over periods of time, noting changes and paying attention to detail.

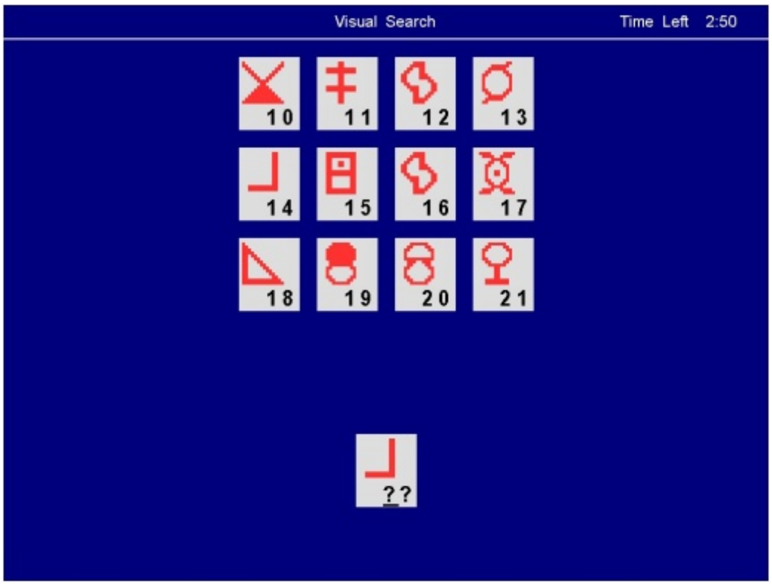
The Vigilance Test (Figure 15) is designed to assess the ability to scan information and switch between tasks. The participant is required to scan pictorial information presented in a matrix format, switch between routine and priority tasks whilst working quickly and accurately.

Figure 15: Vigilance Test



The Visual Search Test (Figure 16) is designed to assess the ability to scan information under time constraints by scanning information presented as letters or line figures, searching for a target among a number of distracters whilst working quickly and accurately.

Figure 16: Visual Search Test



The Table Reading Test (Figure 17) is designed to assess the work rate in terms of scanning and cross-referencing tables of information. The candidate is required to cross-reference row and column numbers to identify values in a table, use multiple tables to identify values when given specific values and interpret the information presented.

The Cognitive Updating Test (

Figure 18) assesses the ability to manage and coordinate tasks in a busy working environment. The participant is required to deal with a series of tasks against the time shown in a digital clock, e.g. monitor, control, set up and adjust a variety of technical systems and update information as required and completed by means of using multifunction displays, interactive pages and a menu system.

Figure 17: Table Reading Test

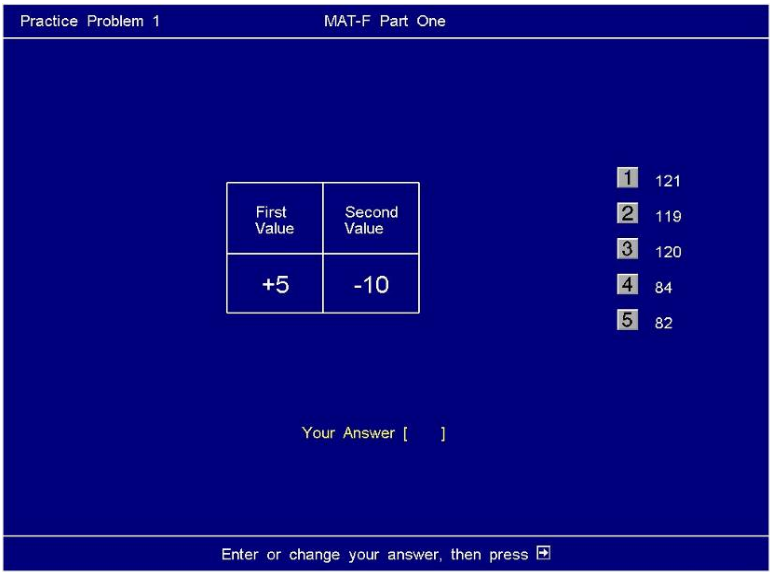
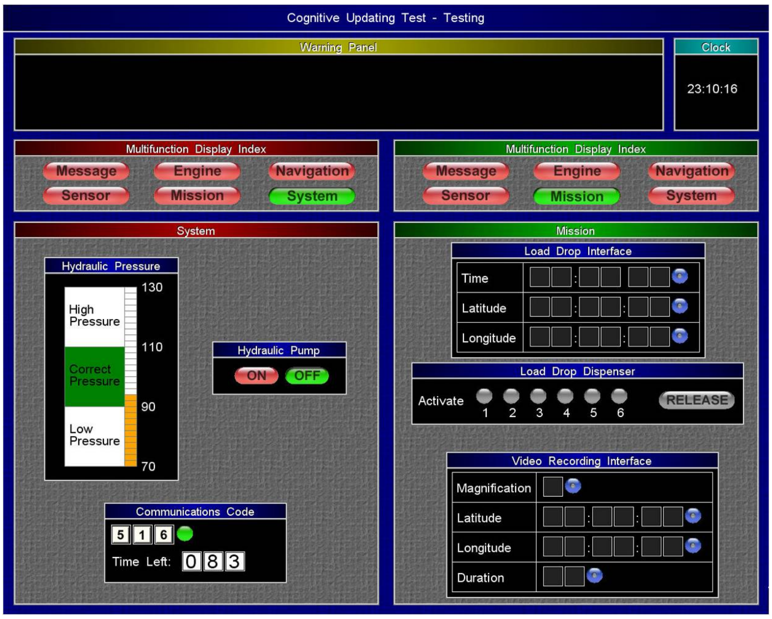
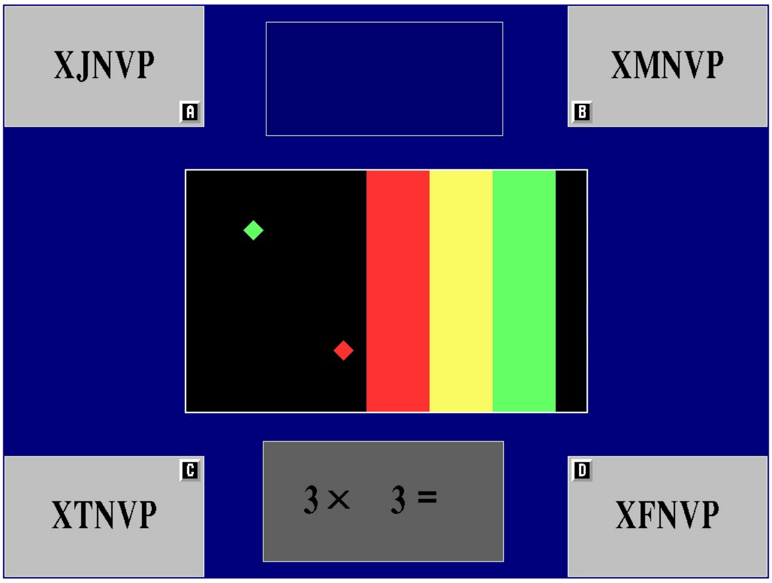


Figure 18: Cognitive Updating Test



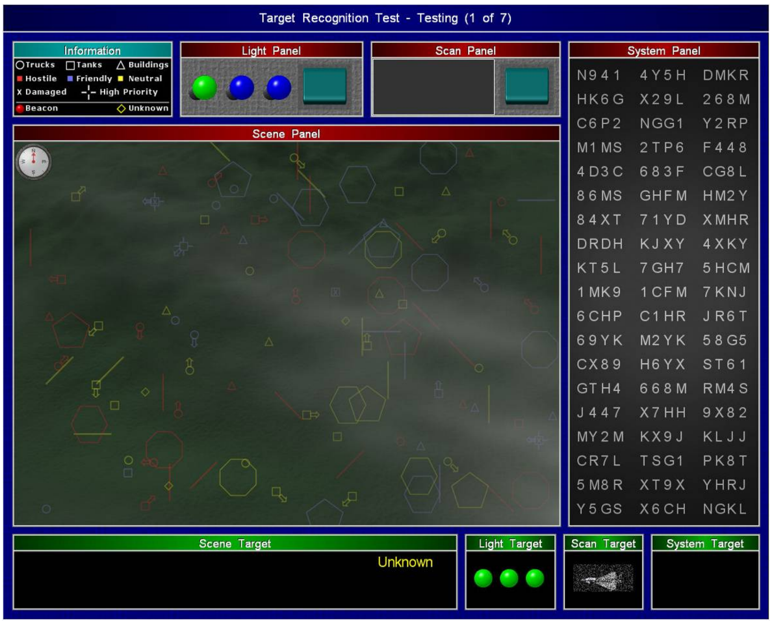
The Colours and Numbers Test (Figure 19) is specifically designed to assess the ability to shift attention between different tasks. The tests require the candidate to deal with more than one task at a time, to monitor changes and pay attention to detail, to remember sequences of information and to complete simple mental arithmetic.

Figure 19: Colours and Numbers Test



The Target Recognition Test (Figure 20) is a multiple tasks test designed to assess the participant's ability to search and identify a series of visual targets. The participant is required to scan and search for a variety of targets which can be imageries, colour patterns, strings of alphanumeric characters, codes/symbols and warning signs whilst prioritising and alternating between the various scan/search tasks and registering as many targets as possible.

Figure 20: Target Recognition Test



Up to 45 candidates are able to undertake the test simultaneously in a purpose-built test room with constant temperature, lighting and sound.

After the completion of the tests programme each candidate is given their test results and debriefed by the Reviewing Officer of the OASC staff. The candidates also have the opportunity to discuss their branch options with their results.

Wanzel (University of Toronto) pointed out parallels between pilot training and surgical training in his paper titled "Teaching the surgical craft: From selection to certification" [15]. Based on how Bell [16] had described in the Annals of the Royal College of Surgeons of England, the battery of tests used by the Royal Air Force that cover the aptitudes that are believed to be relevant to being a pilot seem also to be relevant to surgical proficiency. To test the criterion validity of the Royal Air Force selection test, Bell in his study monitored 200 applicants who were selected for pilot training. All of whom had "probability scores" derived from regression equations from their aptitude test; previous flying experience; age (the younger the higher the score) and motivation to become a pilot. A linear relationship was produced from low probability scores that gave a low probability of success to a plateau at approximately 80% to 85% success at probability scores above 70. At the end of the follow up, 81% of trainees who had scores of 70 and above had passed through flying school and only 22.5% of those who scored below 70 passed.

1.3 Surgical Training & Assessment.

1.3.1 Evolution in Surgical Training

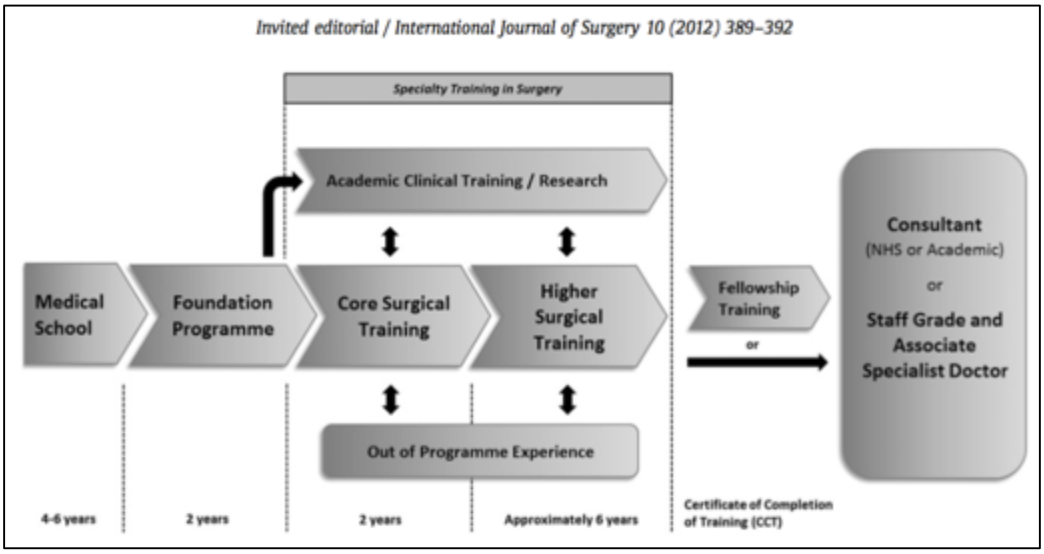
Apprenticeship was the most common and well-established method of surgical training over the centuries, where the student learnt by direct observation and later by imitating the actions of their skilled mentor. The four Royal Colleges of surgeons in the UK have a very rich history dating back half a millennium and have produced many great world-renowned surgeons, promoting the highest standards in surgical education and training.

Dr William Halsted known for being the inventor of the US residency programme, incorporated the bed-side teaching method of his fellow founding professor of John Hopkins Hospital Sir William Osler MD, and the formal training of the German surgeons into producing the Halstedian training model. Dr Halsted was key for the major shift from the apprenticeship model to a more formalised and structured form of postgraduate surgical education that took place at the end of the 19th Century and beginning of the 20th[17].

There have been various changes in postgraduate medical training in the UK over the 20th century, but the greatest reform occurred with the introduction of the Modernising Medical Careers in 2007. After mandatory completion of 2 years of Foundation Training (which was introduced in 2005 to replace the Pre-Registration House Officer (PRHO) and the first year of Senior House

Officer (SHO) years, many specialities adopted a run through programme which later some specialities such as surgery and medicine opted to decouple into core and specialty training. Figure 21 shows the current training pathway published by the UK Association of Surgeons in Training (ASiT)[18].

Figure 21: UK Surgical Training Pathway (2012)[18]



All Specialty selection is centralised and carried out via a national recruitment office. Up to date information for applicants is available in each year's Applicant Handbook[19] via the specialty training website (specialtytraining.hee.nhs.uk). The applications to different specialties at entry levels ST1 (Specialty Training year 1), CT1 (Core Training year 1) or ST3 are led by either Royal Colleges (such as all medicine sub-specialties being led by the Royal College of Physicians) or by various Local Education and Training Boards.

A proposal for a pilot surgical training programme titled "Improving Surgical Training"[20] was published by the Royal College of Surgeons in 2015 based on the recommendation from the Shape of Training review[21]. The independent review looked into whether changes are required in postgraduate medical training to ensure it continues to meet the needs of patients and health services in the future. The pilot curriculum planning to start in 2018 would be a run-through, competence-based programme with a minimum duration of six years to a maximum of eight years[22]. The pilot trial's aims are to improve the quality of training, to better balance training-service for trainees, and look to develop other members of the team from other professional backgrounds to work alongside trainees to improve patient care.

Currently, successful application to Specialty Training needs to follow completion of core surgical training and passing of the MRCS (Membership of the Royal College of Surgeons) examination[23]. The MRCS examination is an intercollegiate examination and it is delivered by all of the four UK Surgical Colleges (Royal College of Surgeons of England, the Royal College of Surgeons of Edinburgh, The Royal College of Physicians and Surgeons of Glasgow and the Royal College of Surgeons of Ireland). The MRCS is comprised of a written part A and an OSCE (Objective Structured Clinical Examination) based part B. Unlike medical specialties, surgical specialties also have an exit exam, which is undertaken during the last two years of training (the FRCS, Fellowship of the Royal College of Surgeons[24]). This too has a written and a clinical component.

The written surgical test requires the candidate to learn large numbers of facts in isolation and Bulstrode and Hall[25] suggest that it might be a surrogate test of motivation to become a surgeon or a rite of passage rather than acquisition of useful surgical knowledge. Adult learning theory suggests that knowledge is best acquired when it is related to the relevant and practical tasks whilst it is being performed. The second part of the exam is more practical and the trainee surgeon will rotate between different skill stations.

Bloom's taxonomy[26] of learning was devised by a committee of educators led by Benjamin Bloom between 1949 and 1953 and it has also been applied to medical education. As seen on Figure 22, the knowledge-based cognitive domain shows how remembering 'sheer volume of facts' is at the bottom level of the taxonomy. Depending on how the written questions in exams such as the surgical intercollegiate examinations have been designed, the level that the trainee surgeon may have to demonstrate is the bottom level, made of remembering facts, terms, basic concepts and answers.

Figure 22: Cognitive Domain: Bloom's Taxonomy of Learning[27].

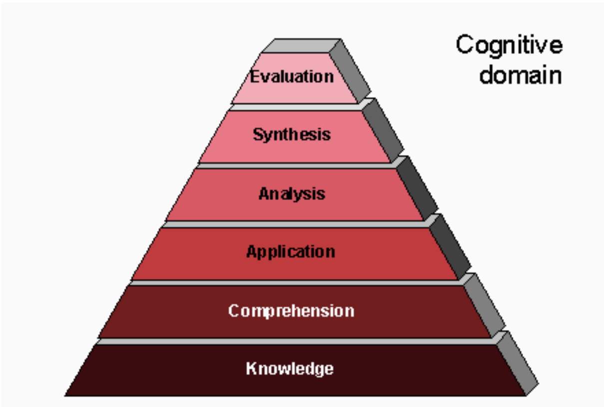


Table 1 was published in the British Journal of Hospital Medicine by Professor Judy McKimm and Tim Swanwick[28] at the London deanery for medical education as a guide for clinical teachers on how to set learning objectives. Listed are the description and useful verbs associated for the levels within Bloom's taxonomy.

Table 1: Writing Objectives in the Cognitive Domain

	Description	Useful verbs for outcome level statements
Knowledge	Recall of information previously presented	Define, list, name, recall, record
Comprehension	Grasping the meaning but not extending it beyond the present situation	Describe, explain, discuss, recognize
Application	Using the rules and principles	Apply, use, demonstrate, illustrate, practice
Analysis	Breaking down components to clarify	Distinguish, analyse, calculate, test, inspect
Synthesis	Arranging and assembling elements into a whole	Design, organize, formulate, propose
Evaluation	Ability to judge X for a purpose	Judge, appraise, evaluate, compare, assess

Part A of the intercollegiate surgical examinations use 'single best answer multiple choice questions' and 'extended matching questions'[29] to assess surgically relevant clinical knowledge. Referencing the above table, such questions may have limitations on assessing beyond the 'knowledge level' of a trainee. Comprehension or application of surgical knowledge may be difficult to demonstrate. Denis Diderot, French philosopher and author and famous for editing the Encyclopédie wrote "It is handicraft that makes the artist, and it is not in books that one can learn to manipulate"[30].

Part B of the intercollegiate exam has been evolving over the years and currently is in Objective Structured Clinical Examination (OSCE) format[29], where the aim is not only to assess the basic surgical scientific knowledge but also its application into clinical surgery in appropriate contexts (third level on the Bloom's taxonomy in Figure 22). OSCE assessments in clinical specialities (not only medicine, but other health professions such as nursing or midwifery) has been introduced over the last forty years[31] and is used in both undergraduate and post graduate clinical assessments.

In the 1975 BMJ article Harden et al[31] described OSCE as an assessment method as a circuit of short stations, where real or simulated patients and one or two examiners (more than one to reduce inter-rater reliability). The design of the OSCE stations are meant to provide an objective assessment where all the candidates have to pass through all the stations, using the same marking scheme; structured stations where there is a very specific task to be achieved with clear written instructions and prior educated simulated patients and examiners; and a clinical examination where clinical as well as theoretical knowledge is tested by the examiner from standardised questions on the mark sheet. The questions during the short OSCE stations could assess higher level of knowledge by demonstrating analysis and application of surgical practice (higher up in the cognitive domain of Bloom's taxonomy).

The 18 OSCE stations at the Intercollegiate Membership Examination Part B is comprised of various stations which include anatomy and surgical pathology (5 stations), applied surgical science and critical care (3 stations), clinical and procedural skills (6 stations) and communication skills (4 stations). Suturing, insertion of a chest drain or urethral catheter, fine needle aspiration or debridement of a wound are examples of the OSCE practical stations. The ability to perform manual tasks related to surgery which demands manual dexterity, hand/eye coordination and visual/spatial awareness is tested in the procedural skills stations which may be fewer than 3 stations [29].

1.3.2 Surgical Selection

The task for the Royal Colleges and the surgical associations is to develop a fair, open and reliable method for selecting surgical trainees. The selection process should define the attributes that make a good surgical trainee which principal objective would be to identify and select those trainees who will eventually become the best specialty consultant surgeon. Gallagher et al[32] from the National Surgical Training Centre at the Royal College of Surgeons in Ireland writes that the selection process should also be able to identify those candidates who are likely to be unsuccessful or problematic as future surgeons and discourage them from pursuing a career in surgery. The cost of selecting the wrong candidate is not only expensive but have detrimental effects on patients, colleagues, hospital, the individual and their families.

The surgical colleges and associations should search for validated tools that exist to measure the identified attributes, and if there are attributes that are not being currently measured, then there need to be development and validation of new tests. To validate the whole selection process, there needs to be careful consideration in weighting each measure used.

Consultant Surgeon and educator Mr P Tansley[33] presented on the visuo-spatial and technical ability in the selection and assessment of higher surgical trainees at the winter meeting of the British Association of Plastic Surgeons at the Royal College of Surgeons of England in December 2005. It was found that during the selection process into higher surgical training found no specialities to have assessed innate visuo-spatial skills or technical abilities. Tansley suggested that assessment should include practices used in aviation training, including the use of high-technology systems. A suitable validated model should be piloted so that trainees are selected appropriately, demonstrate proficiency in clinical skills and most importantly act as an early warning system to identify those who do not meet the required standards. He concluded that not incorporating such process in the 21st century may result in inappropriate selection, inefficient training with wasted resources and loss of public confidence.

1.3.3 Objective Structured Assessment of Technical Skills

Since the national centralisation of recruitment, it has been feasible to better standardise the selection process and to introduce technical skills stations relevant to the speciality and to the level of entry. The practical OSCE stations are marked using the scoring method of Objective Structured Assessment of Technical Skills (OSATS).

The use of OSATS in testing operative technical skills using bench model simulation was described by Reznick et al[34] in 1997. A systematic literature search by Shaharan and Neary[35] (2014) had found the OSATS to be the main scoring system to evaluate surgical candidates' technical skills performances in open and laparoscopic skills.

Prior to the introduction of a technical skills test station at the surgical selection interview process there have been many publications dealing with the importance of having an objective means to assess such technical skills.

The Basic Surgical Skills Course run by the Royal College of Surgeons uses the OSATS for the assessment of the participants (Appendix 1).

Table 2: OSATS sample by Reznick. Am J Surg 1997[34]

STATION3

SMALL BOWEL ANASTOMOSIS

INSTRUCTIONS TO CANDIDATES

You have just resected a segment of small bowel. Perform a single layer. interrupted, end to end anastomosis to restore continuity

ITEM		Not Done or Incorrect	Done Correctly
1.	Bowel oriented mesenteric border to mesenteric border, no twisting	0	1
2.	Stay sutures held with hemostats	0	1
3.	Selects appropriate needle driver (Gen surg, medtip/med or short length)	0	1
4.	Selects appropriate suture (atraumatic, 3.0/4.0, PDS/Dexon/Vicryl/silk)	0	1
5.	Needle loaded 1/2 to 2/3 from tip	0	1
6.	Index finger used to stabilize needle driver	0	1
7.	Needle enters bowel at right angles 80% of bites	0	1
8.	Single attempt at needle passage through bowel 90% of bites.	0	1
9.	Follow through on curve of needle on entrance on 80% of bites	0	1
10.	Follow through on curve of needle on exit on 80% of bites	0	1
11.	Forceps used on seromuscular layer of bowel only majority of time	0	1
12.	Minimal damage with forceps	0	1
13.	Uses forceps to handle needle	0	1
14.	Inverting sutures	0	1
15.	Suture spacing 3 to 5 mm	0	1
16.	Equal bites on each side 80% of bites	0	1
17.	Individual bites each side 90% of bites	0	1
18.	Square knots	0	1
19.	Minimum three throws on knots	0	1
20.	Suture cut to appropriate length (does not interfere with next stitch)	0	1
21.	No mucosal pouting	0	1
22.	Apposition of bowel without excessive tension on sutures.	0	1
MAXIMUM TOTAL SCORE			(22)
TOTAL SCORE			
EXAMINER			

A study on the surgical skills assessment of surgical residents by Reznick et al showed the validity and reliability of the OSATS (Table 2) being similar to those of the more traditional Objective Structured Clinical Examination (OSCE) and could be acceptable for summative high-stakes evaluation purposes[34].

Gallagher et al[32] from the National Surgical Training Centre at the Royal College of Surgeons in Ireland published the assessment method used for the selection of trainees for higher surgical training in general surgery in Ireland in 2006. Traditionally the selection process consisted of traditional methods measuring academic achievement, publications and research. These were used as they are easily measured and believed to be surrogate markers of achievement. The study described candidates who were first shortlisted via a detailed structured application form but then were entered into an objective, structure assessment of surgical technical skills. Each candidate rotated through 10 surgical skills stations (Table 3); the first described was the skills station of a laparoscopic cholecystectomy on a Limbs & Things model where the candidate had to describe each stage of the procedure whilst the examiner acted as a first assistant. The other skills stations were; Lichtenstein hernia repair, arterial closure with Dacron patch, end-to-end bowel anastomosis, resection of ingrown toenail, core laparoscopic skills, sapheno-femoral junction ligation and division, upper and lower gastrointestinal endoscopic examination (in a Simbionix GI endoscopic simulator[36]) and excision of subcutaneous lesion.

The study found that those candidates who performed well at the skills stations also scored high in the rest of the interview process. The objective assessment of surgical skills accounted for only 10% of the overall mark but involved the highest allocation of resources and expenditure. The simulated skills tests used in this study would not be possible to execute for the majority of the current Core Training year 2 (CT2) candidates applying for Specialty Training year 3

(ST3) (equivalent to the first year of higher surgical training) due to the lack of experience in most of the listed tests.

Table 3: Objective Assessment of Surgical Skills[32]

Table 2. The surgical tasks candidates were required to perform at each of the 10 skills stations

Skills station	Candidates' surgical skills assessment task
Laparoscopic cholecystectomy (30 min duration)	To carry out a laparoscopic cholecystectomy (on a Limbs & Things model). They were also required to describe each stage of the procedure. The examiner acted as a first assistant under the instruction of the candidate
Lichtenstein hernia repair (30 min duration)	To carry out a Lichtenstein hernia repair (on a Limbs & Things model) before which they had to describe each step of the procedure
Arterial closure with Dacron patch (30 min duration)	To carry out an arteriotomy (Annex Art) 1.5 cm long and trim the edges and then insert a Dacron patch
End-to-end bowel anastomosis (30 min duration)	To carry out an end-to-end bowel anastomosis (on a Limbs & Things model) using a single layer of interrupted extramucosal sutures
Resection of ingrown toenail (30 min duration)	To identify (on a Limbs & Things model) the growth plate, the blood supply, the digital nerve and the location of the anaesthetic ring block before resection of the ingrown toenail (on a Limbs & Things model)
Core laparoscopic skills: object positioning and sharp dissection (20 min duration)	To complete two separate tasks on the Promis simulator (Haptica). The first task involved the bimanual movement of beads to predefined locations and the second involved the excision of a clearly marked triangle from the upper layer of an outstretched rubber glove
Saphenofemoral junction ligation and division (20 min duration)	To carry out dissection of the sapheno femoral junction and vessels; Control all tributaries and ligate and divide the sapheno femoral junction vessels (on a Limbs & Things model).
Upper GI endoscopic examination (20 min duration)	To complete an upper GI endoscopic examination and identified appropriate sedation, including dosage and potential side-effects and also to identify important anatomical structures and landmarks, such as the oesophagogastric junction, antrum, and so on (Simbionix)
Lower GI endoscopic examination (20 min duration)	To carry out a full colonoscopy and identify any pathology present and then repeat the colonoscopy and carry out any intervention they felt was indicated, for example, polypectomy (Simbionix)
Excision of subcutaneous lesion (20 min duration)	To complete an excision of a subcutaneous cyst with the appropriate incision and then close the wound with interrupted sutures (on a Limbs & Things model)

GI, gastrointestinal.

Information on the selection process was available for Cardiothoracic surgery at both ST1 and ST3[36] levels online using Objective Structured Assessment of Technical Skills (OSATS) scoring method (Table 4). The interview for ST3 also included six structured questions, a communication station with an actor and a portfolio station.

The selection interview for the national recruitment of neurosurgical trainees at both ST1 and ST3 held in Sheffield in 2011 introduced the use of three clinical skills stations, which included suturing; checking tissue handling, the use of microscope to move beads for dexterity; and image guided intracranial instrument insertion for three-dimensional spatial awareness. These three skills stations required a lot of both candidates' and examiners' time and resources but only contributed to 5% to the overall interview marking score. Not only are these tests very labour intensive but the validity of these tests lacks the robustness of longitudinal data.

Table 4: OSATS National Cardiothoracic Selection Interview Scoring for ST1 & ST3[36]

ST1

Knot Tying

Time & Motion
Suture Handling
Flow of task
Knowledge of task
Overall Performance
Quality of final product

Video Assisted Thoracoscopy

Time & Motion
Instrument Handling
Flow of procedure
Knowledge of Procedure
Overall Performance
Task Completion

Suture Station

Suture Station
Needle Grasping
Instrument Handling
Suture tying
Overall Performance
Quality of final product

ST3

Knot Tying

Time & Motion
Suture Handling
Flow of task
Knowledge of task
Overall Performance
Quality of final product

Video assisted Thoracoscopy

Time & Motion
Instrument Handling
Flow of procedure
Knowledge of Procedure
Overall Performance
Quality of final product
Task Completion
Staple Use

Anastomosis station

Arteriotomy
Graft Orientation
Bite appropriate
Spacing appropriate
Use of needle holder
Use of forceps
Needle angles
Needle transfer
Suture management/tension
Knot tying

1.4 Minimally Invasive Surgery

1.4.1 Laparoscopic Surgery

There has been a significant advancement of surgery since the introduction of Minimal Access Surgery (MAS) in the 1980s. Human factors such as psychomotor, cognitive and perceptual faculties have been given more explicit attention. It has been reported that compared to open surgery, the individual difference is greater in laparoscopic surgery[37].

There are various reported factors that contribute to the difficulties in laparoscopic surgery. The tactile feedback is reduced as the tissues are manipulated with long laparoscopic instruments compared to the direct palpation of tissues in open surgery. The fulcrum effect created by the patient's abdominal wall makes the working end of the laparoscopic instrument move in the opposite direction of the surgeon's hand. The majority of the current screen monitors used in laparoscopic surgery display images in a two-dimensional format. Reinhardt-Rutland and Gallagher used the term 'binocularity' in 1996[38] to describe depth perception, which is lost during the visualisation of laparoscopic surgery, and later Gallagher described this again as pictorial perception in 2003.

At the Second International Conference on Surgical Education and Training (ICOSET), Steven Dawson from the Harvard Medical School presented on the use of simulation on training as well as assessment. He stated that the famous medical saying of “first do no harm” is very apt on the use of simulation for surgical training. Surgery will join the other high-risk disciplines like aviation in the use of simulated experiences to make actual experiences safer. He finishes with the following sentence: “When surgical performance assessments methods become robust we can end the worst tradition that still preserve, i.e. training the least skilled amongst us by using the most gravely ill members of our human family as teaching material”[39].

A systematic review of skills transfer after surgical simulation training[40] of ten randomised controlled trials and one non-randomised comparative study showed that skills acquired by simulation-based training seemed to be transferable to the operating setting. A more recent 2014 systematic review on skills transfer after laparoscopic cholecystectomy and endoscopy simulation based training[41] also showed that skills acquired by simulation based training seem to be transferable to the operative setting. This latter study reviewed seventeen randomised controlled trials and three non-randomised comparative studies and in most cases the simulation based training was in conjunction with clinical patient based teaching programmes.

The level of reality in a simulator is of great importance if it were to provide training in transferable skills to the operating table[42]. A simulator test for dexterity so basic that does not resemble clinical skills will show no correlation such as the one described in a paper by Christopher Tang[43] where surgical trainees were asked to carve soap and this was then correlated with surgical skills.

The augmented reality simulator such as the ProMIS simulator was used by Fabien Leblanc[44] at the Laparoscopic Colectomy Simulation Course at the 2009 American Society of Colorectal Surgery Annual Meeting. The study compared the use of a high fidelity simulator and the human cadaveric model on laparoscopic sigmoid colectomy. As shown in other studies[45], the use of the human cadaveric model resulted in a better overall satisfaction level than the simulator. The use of the human cadaveric model has been known to be the most important factor of course choice for participants as it provides the best anatomic and clinical-like model for surgical procedural training[46]. The use of plastic layers replicating different tissue planes and the use of real surgical instruments on a real physical model still provided with tactile feedback for trainees and standardised laparoscopic sigmoid colectomy techniques could be applied to the simulated model. The initial cost of a simulator might be high but the running costs can be a fraction of a human cadaver and does not require operative facilities.

1.4.2 Basic Surgical Skills

The test chosen for the open surgical skills test in this study is been based on the Intercollegiate Basic Surgical Skills Course (BSS)[47]. The BSS is a two-day course, aimed at junior trainees pursuing a career in surgery who are expected to be able to perform a range of basic surgical skills prior to commencing specialty surgical training. These skills are not specifically covered during their foundation or core training placements. The course aims to teach, assess and certify the ability of trainees to use safe and sound surgical techniques within a controlled workshop environment that are common to all forms of surgery.

The two-day course covers three main areas: Open surgery, electrosurgery and endoscopic surgery. For the purposes of this study only the open surgery tests were used. The learning outcomes of the open surgery course expected the participants to:

- understand and practice safe operating techniques, gowning and gloving;
- understand and practice the correct techniques for laying safe surgical knots;
- understand the characteristics and handling of surgical instruments;
- understand that careful and sound aspects of technique are more important than simple manual dexterity or speed;
- understand the principles of handling tissues and recognise differing requirements for differing sites (e.g. skin, bowel, abdominal wall, vessels and tendons);
- understand and practise local anaesthetic techniques;
- understand the principles of assessment, and primary surgical management of infected and contaminated soft tissues.

The BSS course format includes hands-on practice, individual tuition and assessment (Appendix 1) with personal feedback.

1.5 Aims & Hypotheses of this study

The importance of technical skills in safe surgery is paramount. The use of simulation has increased exponentially over the last decade in both selection and training in undergraduate and postgraduate medical education.

The aim of this study is to find whether the validated computerised Flying Aptitude Test used for the selection of British military pilots can be used to test the technical skills in laparoscopic and open surgical simulation tasks. The introduction of an easily reproducible, portable computerised aptitude test prior to the start of postgraduate medical training could be used to determine technical skills ability and guide the postgraduate trainee into a career in surgery.

Hypotheses:

- The Flying Aptitude Test results will inversely correlate with the total Laparoscopic Simulator tests time.
- The Flying Aptitude Test results will correlate positively with the Basic Surgical Skills tests results.
- Previous surgical experience (undergraduate vs postgraduate trainees) will not show a difference in the Flying Aptitude Test scores.
- There is no difference in gender in the Flying Aptitude Test results and the Surgical Simulation tests results.

Chapter 2: LITERATURE REVIEW

This chapter reviews the literature on the acquisition and assessment of technical skills of surgeons, how the post graduate education system has evolved over time and about the use of simulation in medical and surgical education.

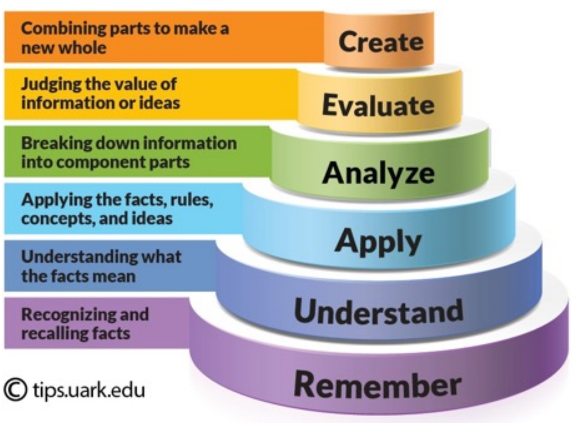
2.1 Acquisition of Technical Skills

Benjamin Bloom was an educational psychologist at the University of Chicago and through a series of conferences during 1953-1959 chaired a group of educators who devised the hierarchical models used to classify educational learning objectives into levels of complexity and specificity to improve communication between educators on the design of curricula and examinations.

His 1956's first volume of "Taxonomy of educational objectives: the classification of educational goals" [48] outlined a classification of learning objectives that has come to be known as Bloom's taxonomy. It remains a foundational and essential element within the educational community and it is also included during the teaching and learning courses in undergraduate and post graduate medicine. Bloom's is hierarchical, meaning that learning at the higher levels is dependent on having attained prerequisite knowledge and skills at lower levels. Updated since the original publication, Bloom's Taxonomy include the following six levels of learning [49] (Figure 23):

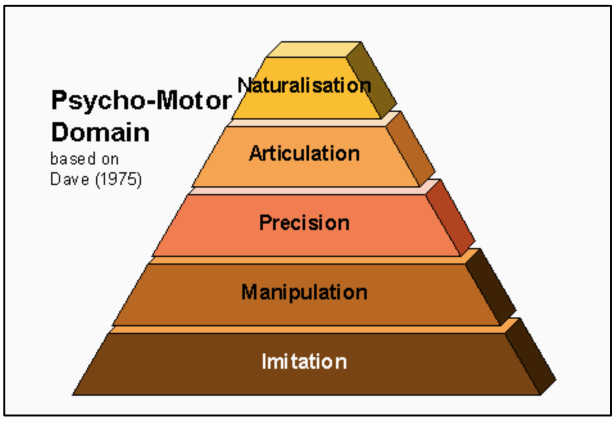
1. **Remembering:** Retrieving, recognising, and recalling relevant knowledge from long-term memory.
2. **Understanding:** Constructing meaning from oral, written, and graphic messages through interpreting, exemplifying, classifying, summarising, inferring, comparing, and explaining.
3. **Applying:** Carrying out or using a procedure through executing, or implementing.
4. **Analysing:** Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose through differentiating, organising, and attributing.
5. **Evaluating:** Making judgments based on criteria and standards through checking and critiquing.
6. **Creating:** Putting elements together to form a coherent or functional whole; reorganising elements into a new pattern or structure through generating, planning, or producing.

Figure 23: Bloom's Taxonomy of Learning [50]



The addition of the Psychomotor Domain into the Bloom's taxonomy of learning was not by Bloom and colleagues but later educators such as EJ Simpson (1972)[51] and AJ Harrow (1972)[52] and RH Dave (1975)[53]. It follows the Bloom's Taxonomy of Learning where technical skills acquisition is organised as a series of levels or pre-requisites, where it is suggested that the next level cannot be addressed until those below have been covered (Figure 24).

Figure 24: Psycho-Motor Domain based on Bloom's Taxonomy of learning.[53][27]



The three progressive phases of learning a new skill proposed by P. M. Fitts and I. M. Posner in 1967[54] is widely accepted in both motor skills literature as well as the surgical literature[34][55]

Prof of Surgery Dr Reznick incorporated the Fitts Posner Three-stage theory of motor skill acquisition into clinical skills acquisition[54] in surgical training (Table 5). The first stage where the learner intellectualises the task (The Cognitive Stage); the second stage where practice and feedback translates conceptual knowledge into appropriate motor behaviour (Integration Stage); and last where the task becomes automatic, speedier, efficient and with precision (The Automation Stage). The learner starts erratic with distinct steps into a more fluid into continuous and adaptive performance.

Table 5: The Fitts Posner Three-Stage Theory of Motor Skill Acquisition[54]

Stage	Goal	Activity	Performance
Cognition	Understand the task	Explanation, demonstration	Erratic, distinct steps
Integration	Comprehend and perform mechanics	Deliberate practice, feedback	More fluid, fewer interruptions
Automation	Perform the task with speed, efficiency, and precision	Automated performance requiring little cognitive input, focus on refining performance	Continuous, fluid, adaptive

Ackerman's Theory of skill acquisition[56] published in 1988 (illustrated in Figure 25) is described as occurring in three qualitatively different phases which is based on the Fitts and Posner (1967) theory; the first cognitive phase is associated with general intelligence and abilities such as verbal, spatial and numerical. The second intermediate phase was highly associated with perceptual-speed abilities and the final phase, where skill is automatized, was associated with psychomotor abilities.

Studies on expertise have been covered numerous times in different fields from medicine to sports but similar themes can be found on how that is achieved.

Regardless of field, Richman[57] wrote that it takes at least 10 years to become an expert and how practice is such an important factor. There are psychologists such as Schneider & Shiffrin[58] who suggest that under certain circumstances extended practice can lead to improved performance which can minimise the range of inter-individual differences. Thorndike described the 'law of exercise' where the stimulus response bond becomes stronger the more it is practiced but only if the feedback is positive the performance is enhanced[59]

Figure 25: Ackerman's Theory of skill acquisition[56]

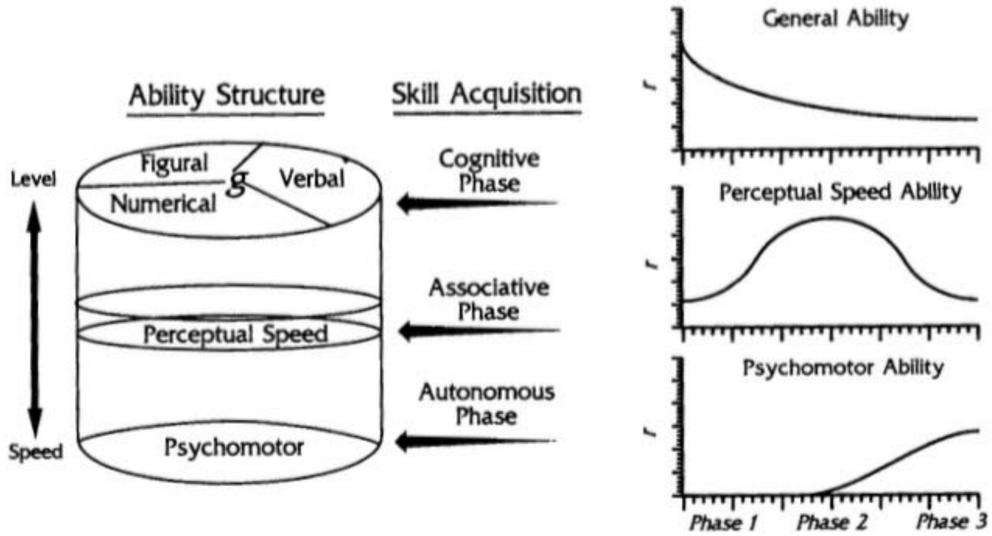


Figure 1. Left: A modified radex-based model of cognitive abilities. Right: Hypothetical ability–performance correlations during three phases of skill acquisition, derived from the framework. The hypothetical task is moderately complex and provides for consistent information processing. From "Determinants of Individual Differences During Skill Acquisition: Cognitive Abilities and Information Processing" by P. L. Ackerman, 1988, *Journal of Experimental Psychology: General, 117,* pp. 291, 294. Copyright 1988 by the American Psychological Association.

Ericsson[60] explains that the role of practice reduces the inter-individual variability when the tasks are within the capabilities of most performers. If the task is not too complex, like driving a car, what limits performance is practice. Individual differences in broad ability measures will show how quickly an individual will grasp the essence of the task and show an initial advantage, but eventually all learners will catch up and minimise the individual differences with time. Some people may take longer than others to learn how to drive but the majority of the population will be able to pass a driving test.

The validity of the ability for measurements to predict individual differences becomes higher when the tasks are highly cognitively demanding, even after extensive practice and long training (Ericsson puts as examples the real jobs of an air traffic controller or a neurosurgeon).

2.2 Assessment of Technical Skills

Capturing performance as it occurs in the “natural” environment has been described as the most useful method of examining expertise by cognitive anthropologists such as Clancey from the NASA/Ames Research Center and Naturalistic Decision Making Researchers such as Ross in the Cambridge Handbook of Expertise and Expert Performance[60]. The online PBA[61] “Procedure Based Assessment” for higher surgical trainees or DOPS[62] “Direct Observational of Procedural Skill” for basic surgical trainees assessment tools are part of the standardised “Work Based Assessment” (WBA) tools for the UK postgraduate surgical trainee. These formative assessment tools are the closest to capturing performance in the “in vivo” environment of an operating theatre but are not used as a summative tool or for selection.

The use of an objective and independent assessment of surgical skills has been important not only for the profession but the public[63] as various publications that suggest poor surgical outcome is the result of suboptimal technical performance in surgery[64].

Melton in 1947[65] described how psychomotor testing required expensive fabrication costs, frequent recalibrations of the equipment, high cost and time of training many examiners to maintain a low examiner: examinee ratio. But with the advent of low-cost general use PCs, the logistics barriers to psychomotor testing started to ease. Validated tests in the military and validated tests for several psychomotor abilities using touch-sensitive computer monitors were produced by researchers such as Ackerman and Cianciolo[66].

The dexterity tests used in this paper were The O'Connor Finger Dexterity and Tweezer Dexterity Tests. These tests had been used in a wide variety of industry to select their workers. They were used to select women for electrical instrument assembly[67] where 85% of the women who scored better than the median on the finger test were successful in electrical instrument assembly. The tests were used to select employees in an electrical fixture and radio assembly[68] watch assembly[69] and electrical shop work[70]. It was also used to predict the success of students in a course in high school shop mechanics[71]. The reliability of the tests were determined by testing 475 men & 215 women from seven occupational groups from unskilled labourers to professionals[10]. Watch factory workers scored higher on the finger and tweezers tests than did general factory workers. The tests selected competent applicants for watch factory work with higher degree of success than in an interview alone[72]. These tests scores were more selective and consistent for success on the job than were age, schooling, previous experience and marital status.

The current UK selection process for surgical training includes practical stations during the interview process. These are simulation stations with a very high examiner demand where at least one or two examiners are required per trainee compared to the use of a computerised system used for skill tests, such as the one used by the military.

The search for a way of selecting trainees who possess the right technical aptitudes was discussed in the 1979 Trans American Ophthalmology Society before the widespread use of minimal access surgery. Thomas J Kirby[73] wrote that some persons are innately gifted or talented with their hands and some are innately clumsy and that most of us fall somewhere between the two extremes. But that an occasional surgical trainee is found to be clumsy with his hands, and becomes a problem and a challenge to train to a level of competence and safety as a surgeon.

The paper had a small number of participants; only 17 surgical trainees were enrolled. Their finger dexterity skills were compared to their teachers' grades. The paper did not show a statistical significant correlation between the two, but the teacher's grades did not use an objective criterion. Even with a statistical non-significant result, the author suggested that the use of such technical aptitude test is of practical value in identifying innately clumsy trainees. Thomas Kirby states that as jobs become less repetitive perceptual and intellectual factors play a more important role. He quotes Anastasi, who comments in Psychological testing[74] that the spatial aptitude or the ability to visualise and manipulate objects in space has become a part of the better tests for general intelligence.

The paper then continues with discussion from other experts in Ophthalmology. Dr Bruce E Spivey asks the question: 'How do we identify a valid and reliable criterion measure for manual dexterity or "good hands"?'. By criterion measure he explains it as some way to test or asses the target performance we are interested in: in this case it might be a direct rating of the surgeon's smoothness or accuracy in a particular situation or it might be an indirect measure such as rate of surgical complications.

He then comments on how dentists test their trainees, saying that they used chalk carving as a measure of perceptual motor ability, ability to follow directions and visualise in three dimensions.

Dr Orbert Machemer explains on how the Storz instrument company were hiring young men with manual dexterity; The applicant has to bring evidence of manual dexterity to the interview (model cars/jewellery) and also perform in a space-relation test (psychological corporation, New York; differential aptitude test-space relations).

He ends his comments stating that most resident and fellows learn to handle surgical instruments by the simple fact of repetition and need not be unusually dexterous. However, given the choice, we would like to find and select the applicant with the most outstanding manual capabilities.

Another ophthalmologist, Professor Loewenstein, in an even earlier 1940 lecture in Prague at the Tennent Institute of Ophthalmology raised similar concerns on the technical aptitudes of surgical trainees. He stated: 'Having seen many young men starting ophthalmology, I am amazed by the surprising fact that our profession (which by general consent requires at least moderate skill and dexterity) appears in Czechoslovakia to attract like a magnet the most adexterous people. One cannot explain this dark riddle of nature, but is it wrong to exclude any person on the grounds that he is handless?'. After discussion, he concluded that ophthalmology must 'examine the beginner, admit the gifted, and refuse the less able'[75].

2.3 Surgical training

Historically the technical competence of surgeons was achieved through many years and long hours under the traditional apprenticeship model. Based on the Socratic teaching method, where inquiry and discussion stimulate critical thinking[76], formal mentoring of surgical trainees at the work place was started by William Stewart Halsted[77] at Johns Hopkins Medical School in the late 1800s and early 1900s and has been a method utilised in many countries until very recently. The way that Halsted mentored his famous trainee and first neurosurgeon Harvey Cushings may have had its roots on Socrates' mentoring of Plato but there have been papers suggesting that this model might not be suited in this new era of reduced working hours, shift working and minimal access surgery[78].

Hall et al, in the British Journal of Surgery paper titled 'Surgeons and cognitive processes'[1] states that the previous operative surgical approach to teaching such as 'see one, do one, teach one' with the emphasis on the number or completed procedures is no longer a feasible one. The reduction in the total number of years in training and the hours per week at the work place has signalled an end to learning by 'service saturation'.

In 2001 the European Union extended the European Working Time Directive (EWTD) to include doctors in training. The EWTD was introduced gradually for junior doctors in August 2004 beginning with the start of shift work and end to the 24 hour on-call system. It was designed to protect the employees across Europe from working excessively long hours. For all junior hospital doctors and dentist, the limit started at 58 hours, then reduced to 56 hours in August 2007 and finally down to 48 hours in 2009[79].

A survey was conducted by J Lowry and J Cripps[80] between December 2004 and January 2005 on the views of trainee surgeons whose training had been affected after the introduction of shift work; A common complaint on those trainees in partial and full shifts was that the theatre training time had diminished, with more than 80% of Senior House Officers and Specialist Registrars reporting a reduction. In this survey one of the trainees made the following comment on the reduction of training: “May God help the next generation of patients, because the next generation of surgeons may well not be able to”.

After the full introduction of the European Working Times Directive (EWTD) leading to shift work and the hours per week down to an average of 48, a nationwide survey in 2009 of 1,600 surgeons in training from all specialties showed that after the introduction of 48 hour work week more than two-thirds of the trainees reported deterioration in their training[81]. Even though the introduction of the EWTD was designed to limit working hours, so to protect the health and safety of employees, 67% of the responding trainees were attending clinical work while officially off-duty to protect their training and gain further surgical experience, and an even higher number (84%) were working in excess of their allocated hours to maintaining the quality of the service provided and ensure patient care was unaffected. The shift work and extra and unpaid work for both training and service resulted in 86% of surgical trainees to believe that there had been a deterioration in their work-life balance.

The struggle that the surgical community has had in opposing the introduction and the continuation of the EWTD in junior doctors training is that the rest of the medical specialities and surgical specialities have not been holding a uniform opinion. Such confrontation occurred in the letter by the Association of Surgeons in Training (ASiT) in response to the House of Commons discussion held with the British Medical Association (BMA) on the 18th of January 2012[82]. ASiT continued to express that the introduction of the EWTD had a negative impact on surgical training and more importantly, patient safety[83], as surgery being a craft specialty, it is different to others and necessitates clinical time to advance training[84]. The results of a cross-speciality systematic review of postgraduate medical education and patient outcomes following the implementation of the EWTD has shown the reduction in working hours to be inconclusive. This contrasts to the widely documented negative effects it has had on surgical training.

The overall reduction in the hours at work have been shown to affect surgical training, but there has been another major contributing factor to the decreased number of operating opportunities by trainees from having to do 13 hour night shift work[85]. The National Confidential Enquiry into Patient Outcome and Death (NCEPOD)[86] identified a considerable deficiency in the care of emergency patients. The implementation of the guidelines resulted in the reduction of non-urgent operating during the night hours. The reduction in operations performed after midnight fell from 30% in 1986 down to 4% in 2001-2, which on a night shift pattern, the trainee has been left with fewer operative opportunities.

The 2012 Report for the GMC by Durham University titled “The Impact of the Working Time Regulations on Medical Education and Training: Literature Review” highlighted the impact of shift work on medical training[87]. The greatest damage being on craft specialities such as surgery and specialities with high emergency and out of hours workloads. The increasing emergency care workload exacerbated the loss of elective training opportunities at all levels.

The challenges with surgical training have been experienced not only by UK trainees but also by surgical trainees in Ireland and the United States. The implementation of shift work in the Republic of Ireland has resulted in surgical trainees reporting both reduction in the quality of their training as well as reduction in their development of their operative skills[88]. Lonergan et at from the Royal College of Surgeons of Ireland[89] described the increasing challenges in surgical training being due to the greater number of trainees, the changing work practices and the restrictions in working hours.

Even before its full implementation in the UK, the European Association of Neurosurgical Societies complied a paper[90] in 2006 with the inclusion of other surgical disciplines (general surgery, orthopaedic surgery, paediatric surgery, cardio-thoracic surgery, vascular surgery and otorhinolaryngology amongst others) on the negative effects of the reduced working hours on trainees they had experienced over the previous 3 years. Their common Position Paper was sent to the Committee on Employment and Social Affairs in the European Parliament highlight the far reaching and serious consequences on the training of surgical specialists as well as on patient care. They concluded that a reduction of about 30~35% of clinical and operative experience during training was unworkable. The potential impact of the political changes within Europe in surgical training is yet unknown, but the 2014 report by the Royal College of Surgeons state that without returning to the excessive hours of the past, a greater flexibility for training hours will be of benefit [91].

The extent of the reduction in hours was smaller in the US but in 2003 they also experienced the introduction of the mandatory 80 hour work week restriction. The paper by Bell et at[92] in 2009 showed a gap in expectation and operative experience for those residents in the US general surgery programmes. The overall effect of the reduction in the total number of postgraduate surgical training years and the hours in the working week has resulted in the surgical trainee being trained in 6,000 hours compared to the 30,000 of previous generations[93].

Geoff Norman, Professor Emeritus from McMaster University, Hamilton Canada, on his paper titled “Medical education: past, present and future”[94] quoted William Osler, Canadian Clinical Educator and founding professor of John Hopkins Hospital, to be the ‘inventor’ of both the medical residency and the clinical clerkship who said: ‘He who studies medicine without books sails an uncharted sea, but he who studies medicine without patients does not go to sea at all.’ Osler was renowned for his approach to bedside teaching, and his insistence that students learn from their patients.

Norman was concerned not only with the reduction in working hours mentioned above on the impact on medical education, but the reduction in the number of inpatients whose stay is shorter. The modern patients are older, have many co-morbidities and are sicker (Table 6) which has an impact on their suitability for ward learning. The demands on these reduced number of available patients for clinical education became more challenging with the increase in the number of learners (great expansion on the number of medical school numbers over the last 15 years) and competition with other health care professionals in training.

Like many researchers in the field of education, Norman [94] writes how, in addition to learning in the clinical environment, a carefully engineered simulated setting can deliver both undergraduate and postgraduate education. Even though it won't be able to replace clinical experience, studies[95] have shown that even the low fidelity low cost simulation (€10 vs €100,000) can provide as good a learning experience as the high fidelity simulation.

Table 6:NHS comparing 2010 to 2000[96]

• Shorter hospital stay for admitted patients (England): o 6.0 to 4.3 days 2000–2010
• Reduced suitability of patients for learning o Elderly, chronic disease, multi-system o Increase of admissions for patients over 75
• Reduced number of admissions overall o More patients handled on an outpatient basis
• Reduced suitability of ward for learning o More homogeneous, more procedure-orientation

In surgical education, to use simulated tests as a valid measure of operative ability, the simulator performance needs to correlate with actual performance in the operating theatre (Concurrent Validity). Senior surgeons should perform better than junior less experienced trainees, and the simulated tests need to be able to discriminate those who are novices to those who are experts (Construct Validity).

A study by Hove et at (BMJ 2010)[97] evaluated through a systematic Medline search, 104 studies which assessed surgical technical skills. The paper used the Oxford Centre for Evidence-based Medicine levels of evidence for diagnostic studies to determine the quality of the studies reviewed. The level of evidence was regarded high if blinded reference standards were used where the observer was blinded to the training level of the subject and when consecutive participants were used as cohorts for the study. This study concluded that most methods of skills assessments are valid as a formative tool such as for feedback or measuring training progress, but few can be used as a summative form of examination or credentialing.

Other levels of validation included those where the rating was performed by un-blinded individuals. This was described as a less objective method where the reliability as well as validity are affected by the outcome of a test.

If the group of participants were non-consecutive, this can result in bias, as those who have not made it into the study could have been eligible and excluded deliberately.

2.4 The Flying Aptitude Test Domains in Surgery

Psychomotor is the ability to coordinate physical movement to pursuit a task. The use of psychomotor tests in surgery is not unusual. Grober et al[98] published a study on the use of Hand-motion analysis as an intra-operative assessment tool. The study evaluated the intra-operative hand movement frequency and the hand travel distance on live patient urology surgeries in both novice and experienced surgeons. The blinded, case matched assessments of technical skills through unedited surgical videos showed meaningful improvement in the novice surgeon and could represent a feasible, objective and valid measure of technical skills.

Visual Reasoning is the ability to use and interpret verbal or numerical information received via visual or auditory input. Our study did not include the test for numerical reasoning where interpretation of numerical information in the form of fractions, decimals and formulae or in graphs and tables is required as the job analysis did not identify it as a required aptitude. According to Fitts Posner[54] and later supported by Ackerman[99], Verbal Reasoning are part of the initial cognitive phase of skill acquisition and measure of general intelligence and differ from measurements of the psychomotor aptitude where skill becomes automated and more fluid.

Spatial Reasoning is the ability to manipulate diagrammatic or pictorial information presented in a two dimensional form into a three-dimensional representation in the mind. This image manipulation requires visualisation, rotation and orientation of the information presented. Apart from the Image manipulation ability, this aptitude also includes spatial integration where more than one image can be combined in the mind and also add a fourth dimension and predict a trajectory of such an object (the dynamic ability).

A Lancet paper by Wanzel et al[100] found that residents with higher visual-spatial scores in the form-board test and the mental-rotations test scored significantly better in their ability to complete and learn a spatially-complex surgical procedure (a Z-plasty) than did those with lower scores.

Attentional Capability aptitude facilitates multi-tasking (switching attention between tasks whilst dealing with auditory and visual input) and storing medium to long term memory and concentrating over period of times whilst paying attention to detail.

Mergel[101] explains the “Three-Stage Information Processing Model” where input first enters a sensory register, then is processed in short-term memory, and then is transferred to long-term memory for storage and retrieval. The Sensory Register is when input is received from the senses (e.g. visual or auditory) which only lasts a few seconds and disappears unless it is of importance and is transferred into Short-Term Memory, where it can be stored up to 20 seconds or more if rehearsed repeatedly. If information is grouped into meaningful parts the items that can be stored in the Short-Term Memory can be increased.

When information is stored from the Short-Term Memory for long term use it goes into Long-Term Memory and storage. Anderson[102] (1983) writes that repetitive information in Short-Term Memory affects its storage in Long Term Memory. Rote memorisation can force information into Long Term Memory, but if old and new information is linked (deeper levels of processing than just rote memorisation), better retention of material can be achieved.

2.5 Medical Education Research

A recent systematic review and meta-analysis on the use of virtual reality training in laparoscopic surgery by Alaker et al. [103] included 31 studies of which twenty four were randomised controlled trials. The number of participants in the studies ranged from 10 to 228 with mean of 41. Only two studies had participant numbers greater than 65 which explains the small median of 28 participants. The two studies with the larger number of participants were Gallagher's study titled "Prospective, Randomized Assessment of Transfer of Training and Transfer Effectiveness Ratio of Virtual Reality Simulation Training for Laparoscopic Skill Acquisition". The study included 225 participants of which 30 were novices and 195, experts. The study by Tan et al from New south Wales, Australia, recruited 228 participants to the cross over trial examining low versus high-fidelity simulation in basic laparoscopic skills training. The participants ranged from medical student, resident medical officer, intern, registrar and experienced surgeon.

The four laparoscopic box trainer tests used in this thesis have been used by our department in five previously published studies (Table 7) with a mean participant number of 36 and median of 37.

Table 7: Studies using the four laparoscopic box trainer tests

	STUDY	n
1	Assessing Laparoscopic Skills In The Novice [3]	24
2	Laparoscopic Performance In Harrier Pilots Versus Medical Students [2]	42
3	Training Environment Affecting Laparoscopic Simulator Performance [42]	54
4	Harrier Pilots Versus Medical Students In Single Incision Laparoscopic Surgery [104]	37
5	Single Incision Versus Conventional Laparoscopic Techniques [105]	22

Like many medical educational research this is a cross-sectional study and the number of participants was dependant on who was willing to answer to the adverts and participate. The majority of published studies in surgical simulation, including our previously published studies, have relatively small sample sizes.

Chapter 3: METHODS

This chapter describes the demographic data, the recruitment and ethical approval processes followed by detailed descriptions of the methods used for the Flying Aptitude Test and Laparoscopic and Basic surgical simulated tests.

3.1 Demographic Data

There were 243 participants in this study, of which 230 (94.7%) took part in the Flying Aptitude Test at the OASC, RAF Cranwell.

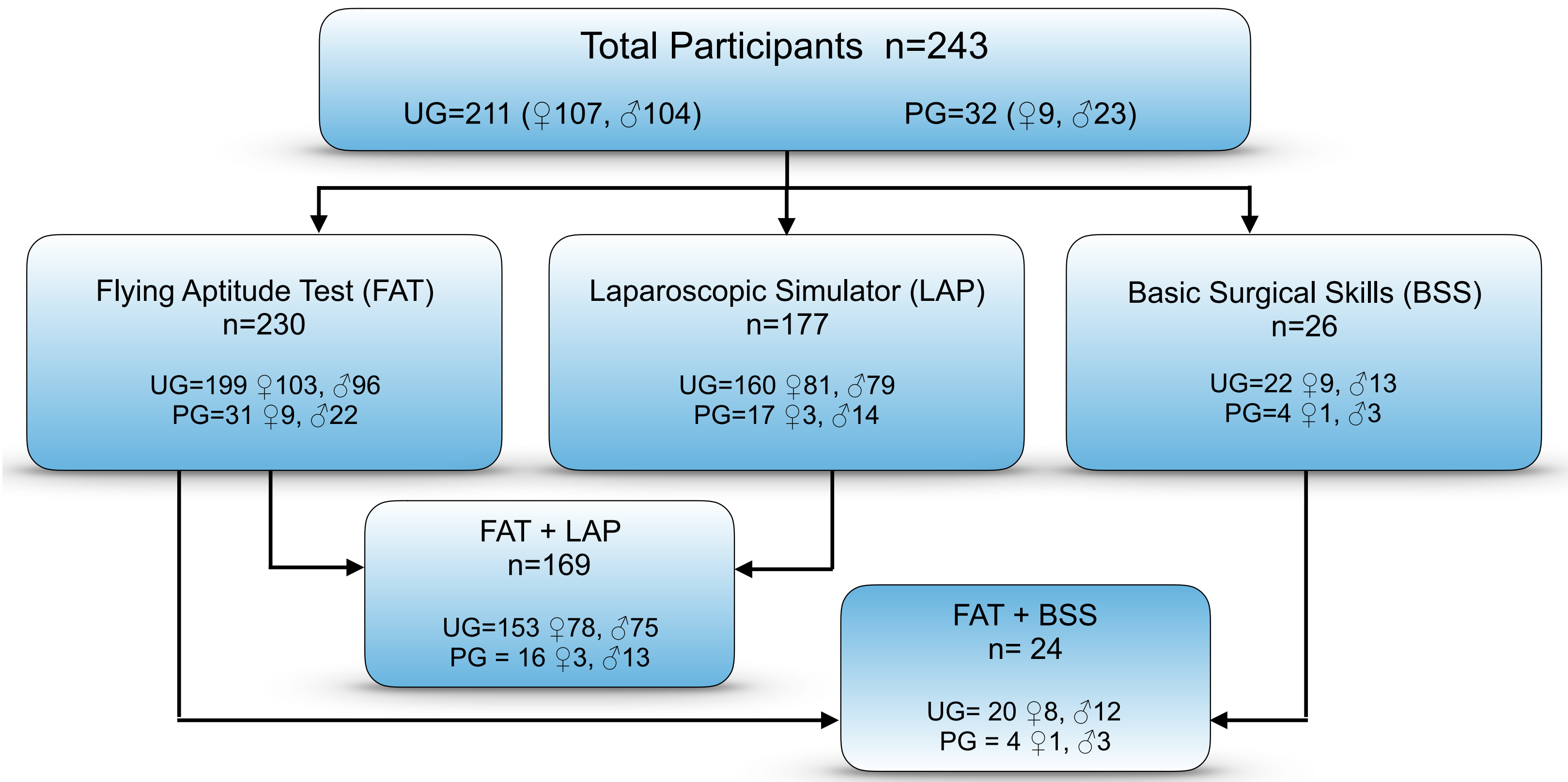
177 participated in the laparoscopic simulated tests and completed the demographic questionnaire (Appendix 6) of which 169 (95.5%) also participated in the Flying Aptitude Test at RAF Cranwell.

26 participants participated in the basic surgical skills test in the same format as it is run by the Royal College of Surgeons as a requisite course during Core Surgical Training[47]. Apart from two participants, they (n=24) also underwent the Flying Aptitude Test.

UG: undergraduates

PG: postgraduates

Figure 26: Participants numbers and tests undertaken



3.2 Ethical Approval

National Health Service (NHS) Research and Development sponsorship was granted by The Research and Innovation Nottingham Integrated Clinical Research Centre (Research and Development at the University of Nottingham Hospitals NHS Trust ref. 10GS002) and the Centre of Aviation Medicine and OASC granted sponsorship for the use of the aptitude test facilities.

The documents submitted for Ethical approval included The Study Protocol (Appendix 2), Consent form (Appendix 5) and Participant's questionnaire (Appendix 6).

The MoD ethical approval interview took place at the Centre of Aviation Medicine and ethical approval for the use of the Flying Aptitude Test was granted by the Ministry of Defence Research Ethics Committee MoDREC (reference 1002/299) (Appendix 7).

The NHS ethical approval was submitted to the North Nottinghamshire Local Research Ethics Committee using the Integrated Research application system. The Derbyshire Research Ethics Committee granted the study NHS Ethical Approval (ref. 10/H0401/43) (Appendix 8).

3.3 Sample Size & Recruitment

The sampling procedure was non-random (nonprobability) as only those participants who self-referred were included. Within the nonprobability sampling methods, the selection of participants was also of convenience sampling. This sampling procedure may contain sources of bias, however the majority of medical education research studies are based on nonprobability sampling, as random (probability) sampling is not only time-consuming and expensive, but also not feasible in some situations[106]. In this study, the maximum sample size was set at 450 participants as it was the potential number of trainees available in the East Midlands deanery and of undergraduate medical students.

There were four groups of participants recruited in this study; The first group included the Nottingham University medical students. They were recruited via internal email, which briefly described the study and were asked to reply for further information. Posters (Appendix 3) were put in the medical school and at the main teaching hospital. Lecture presentations were given at the end of their teaching day for each year group.

The second group included junior trainees after graduation in foundation training and all specialty (medical and surgical) core training. Initially only the East Midlands Deanery trainees were involved, and their emails obtained via the deanery, but an ethical amendment allowed trainees from all deaneries. Presentations were given during deanery teaching days.

The third group included senior trainees in all specialties. They also were recruited via email and lecture presentations.

The last group were the military doctors in training, recruited with the help of the chief of RAF's Centre for Aviation Medicine. Methods included emails and presentation at their annual meeting at RAF Henlow.

Once the participant showed interest in participating in the study, they were sent an email containing the Participant's Information Sheet (Appendix 4) and Consent Form (Appendix 5) with dates to choose for the visit to RAF Cranwell and the visit to the Trent Simulation and Clinical Skills Centre at the Post Graduate Building in the Queens Medical Centre Campus, Nottingham University Hospitals NHS Trust. All participants were self-selected groups.

3.4 Consent

During the recruitment process the participant was sent via email the consent form again at least 24 hour prior to their test in a pdf format (Appendix 4). The participant was asked to then bring the signed consent form to RAF Cranwell to be checked by the principal investigator and countersigned prior to the start of the aptitude test. The consent forms were then signed by the chief investigator.

The last box on the consent form was an optional point where the participant was asked whether they were willing to be contacted in the future for the longitudinal part of a future study, following them up their chosen career path.

3.5 Questionnaire

All participants were asked to complete a questionnaire (Appendix 6) either during their visit to the OASC or whilst attending the Trent Simulation Centre for the Surgical Simulation Tests. Data collected included:

- Demographic: age, nationality, gender, training grade, specialty
- Interest in a career in surgery or gastroenterology
- Previous experience in minimally invasive surgery/scoping
- Computer/video game playing

The participants were asked to rate themselves on their level of computer game skills, from beginner to expert. The ratings will depend on their subjective feelings about their level of skill and be prone to measurement error.

3.6 Flying Aptitude Test Methods

At the Royal Air Force OASC at RAF Cranwell samples of each of the six aptitude domain tests were trialled under the guidance of the OASC chief psychologist and programme designers. The computer programmers were involved by providing the explanation of how to perform each test and chief psychologist from the aptitude design team explained how each aptitude test related to the skills required in the jobs there were selecting their candidates into (detailed explanation of each test in Chapter 1.2.2).

Laparoscopic cholecystectomy (gall bladder removal by minimally invasive surgery) was chosen as the surgical procedure to be analysed at the OASC as this is the commonest type of laparoscopic procedure undertaken in the UK since the first laparoscopic cholecystectomy was performed by Prof Muhe of Boblingen, Germany in 1985[107]. 76,497 cholecystectomies were performed in England in the year 2013/2014[108], of which more than 90% are laparoscopic[109]. Laparoscopic cholecystectomies are part of the curriculum to all surgical trainees within the general surgical programme regardless of subspecialty. For the Certificate of Completion of Surgical Training (CCT) in the UK the Intercollegiate Surgical Curriculum Programme states that every general surgical trainee needs to have a minimum of 50 laparoscopic cholecystectomies recorded as the main surgeon and shown competence through the Procedure Based Assessment (PBA) tool[61]. Laparoscopic cholecystectomy simulation in a box trainer or virtual simulation programme is one of the commonest laparoscopic procedures used in studies to test surgical skills[41,110–113] as well as used for surgical skill evaluation in vivo[114].

A job analysis was performed with the OASC's chief psychologist on the laparoscopic cholecystectomy. The analysis involved deconstructing the surgical procedure step by step from knife to skin as per the ISCP (Intercollegiate Surgical Curriculum Programme) Procedural Based Assessment (PBA)[115] for laparoscopic cholecystectomy. Each point from the skin incision to all the: global and task-specific items within the intra-operative technique up to skin closure were assigned an aptitude following the OASC's standard job analysis method.

The result of the job analysis resulted in the Surgeon's profile to be divided into five different aptitudes; Verbal Reasoning, Short Term Memory & Capacity, Spatial Reasoning, Psychomotor & Attentional Capability. The Numerical and Symbolic Reasoning were omitted as these were not identified as relevant during the job analysis of the laparoscopic cholecystectomy.

Table 8: Sample candidate's Flying Aptitude Test result at the OASC RAF Cranwell

APTITUDE DOMAIN	Domain Weight	Stanine Score
Psychomotor	40%	3
Verbal Reasoning	20%	8
Attentional Capability	10%	2
Spatial Reasoning	20%	6
Short Term Memory	10%	6
Index	100%	53%

The weighting of the five aptitudes was decided depending on the frequency of each aptitude used during the job analysis. Verbal Reasoning 20%, Attentional

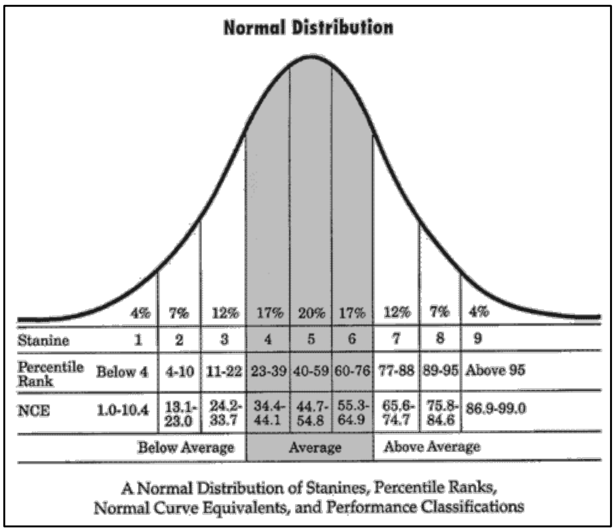
Capability 10%, Spatial Reasoning 20%, Psychomotor 40% and Short Term Memory 10%.

Table 8 shows an example of one of the candidates Flying Aptitude Test results sheet. The first column contains the five aptitude domains, which each was given a weighting out of a 100, psychomotor having the most weight with 40%. The five separate aptitudes results were scored by the OASC in a Stanine scale (STAndard NINE point scale).

The Stanine scoring system[116] is a method of scaling test scores on a nine-point standard scale with a mean of five and a standard deviation of two. The raw scores are converted so the shape of the results will assume a normal distribution.

Each participant's raw score is added into the pool of all the other participants who have undertaken the Flying Aptitude Test at RAF Cranwell over the previous one year, which is mainly made of candidates who wish to join the RAF. The raw scores are rank ordered to make the conversion. These ordered scores are then divided into percentages. The top 4% of the scores will be given a stanine score of 9, the next highest 7% given a score of 8, the next 12% a stanine of 7, the next 17% a stanine of 6. The middle 20% are stanine 5 (which is always the mean). Stanine 1 corresponds to the bottom 4%, stanine 2 to the next lowest 7%, stanine 3 the next lowest 12% and stanine 4 to the next 17% (Figure 27). If our candidates score a 5 at one of the aptitudes then they will be deemed to be average compared to the rest of the RAF candidates, if the score is between 1 to 3 then they are below average and if the score is between 7 and 9 then they are above average.

Figure 27: STAndard NINE scoring system[117]



These Stanine scores were then multiplied by the domain weight and given an Index score, becoming also a percentage as the total adds up to a 100.

3.6.1 Flying Aptitude Test at Officers & Flying Aptitude Test RAF Cranwell

When the participants arrived at RAF Cranwell, a member of staff at the OASC gave them an introductory lecture, which is the same one as any other military candidate will receive when attending the aptitude test. They were introduced to the various tools that they will be using (last picture from Figure 1 page 15) such as the computer Screen, a custom-made keyboard suitable for both right and left handed participants, a joystick, a trackball, a pair of headphones to listen to instructions and pedals.

To minimise the Practice Effects RAF candidates are not allowed to re-take the RAF aptitude test within twelve months and this was emphasised prior to the start of the test in case there were any participants in this study who had an interest in joining the RAF within the next twelve months.

The number of participants per visit did not go over the maximum number of 45 candidates who are able to undertake the test simultaneously in the purpose-built test room so all participants had the same starting time. The individual tests had set times, but the introduction and questions prior to starting the tests could be read at individual pace. Therefore, the finishing time for the participants showed a variation.

After the completion of the tests programme each candidate was given their test results and debriefed by the Reviewing Officer of the OASC staff.

3.6.2 Self-Assessment Form

Prior to the start of the Flying Aptitude Test each participant was given a Self-Rating Form (Appendix 9). The form asked each participant to put a cross along an unmarked line. The question was: ‘How good do you think you will be?’ and the line spread from ‘very bad’ to ‘very well’, which was converted into a percentage value for analysis. This analysis was to see if there was a correlation between the results of the Flying Aptitude Test and the participant’s expectation.

3.6.3 Computerised Flying Aptitude Test

The total time for the six tests was 90 minutes. The tests were timed, but the instructions could be read at each participants' own pace, which resulted in the participants in the group finishing in slightly different times. After the last participant finished all the data was fed to the psychologist team who printed out each person's results and distributed them. The rules of the OASC stated that the participants were not allowed to repeat the test, for example to join the RAF, within 12 months. This is due to the risk of the "Practice Effect", where participants may perform differently in the second condition because of familiarity with the experimental situation. This information was clearly stated in the participant's information sheet prior to joining our study and repeated verbally on the day of the study prior to taking the test.

3.6.4 Feedback Form

After the end of the Flying Aptitude Test and whilst waiting for their results, each participant was asked to fill in an anonymous feedback form (Appendix 10). There were questions on the organisation of the study, on the trip to RAF Cranwell but the question which has been included in this study is:

Q9. Selection into UK Surgical Training may benefit with the introduction of an aptitude test as an addition to the already existing Interview process.

3.7 Laparoscopic box trainer simulator tests Methods

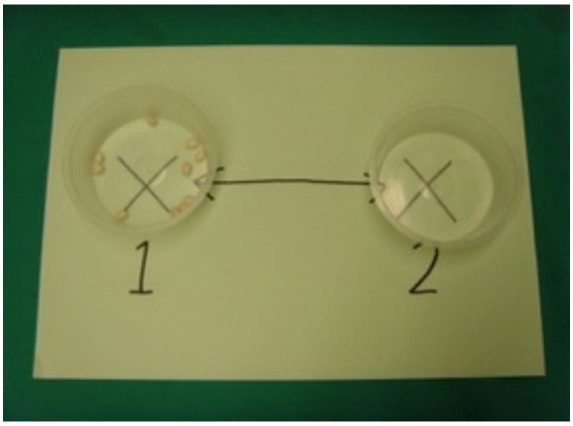
The recruitment for the laparoscopic simulator tests at the Trent Simulation Centre could be done through an online recruitment site (Appendix 11) but the majority of the participants booked their tests directly via email.

A 30 degree laparoscopic camera was used with a box trainer. Four tests were used in this study. All four tests were used in various previous published studies[2,3,42,104,105] which have proven to differentiate between novice and expert laparoscopic surgeons[3].

Each task was explained and demonstrated once prior to start. The two types of instruments (the laparoscopic grasper used in the first three tests, and the laparoscopic scissors used in the last appendicectomy task) were given to the participant so they could familiarise themselves. The method of use was explained and demonstrated and were asked to copy to check that each participant understood the task and use of the instruments. The four tasks were performed standing up in front of the box trainer and screen. The Investigator doubled up as the scope holder standing at the participant's right hand side together facing the screen.

3.7.1 Laparoscopic Test 1: The bean drop

Figure 28: The Bean Drop



First the use of the two graspers were demonstrated to the participant, the task was explained where two plastic see through pots were attached to the floor of the box trainer, with 15 beans in the left pot.

The task is to move ten beans, one by one, to the right cup. If a bean is dropped outside of the pots it counts as an error. The dropped bean should be left and a new bean picked up.

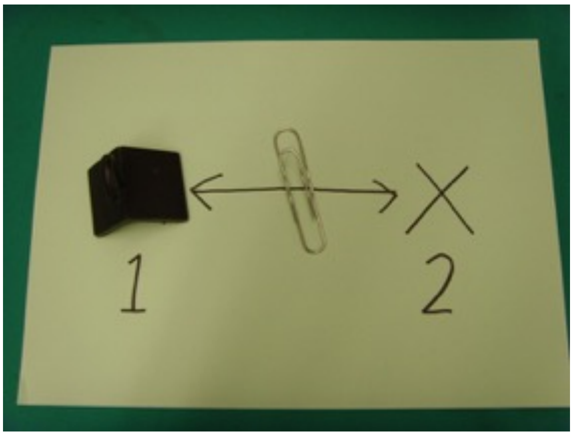
The participants needs to have both left and right graspers inside the box trainer, but is allowed to use whichever one of the two to move the beans one by one. Both pots being in reach of both graspers.

It is explained that time for speed as well as errors for accuracy are measured.

Time starts when both graspers are inside the box trainer and in view on the screen. The time stops when the tenth bean is dropped into the right pot.

3.7.2. Laparoscopic Test 2: The Block Move

Figure 29: The Block Move



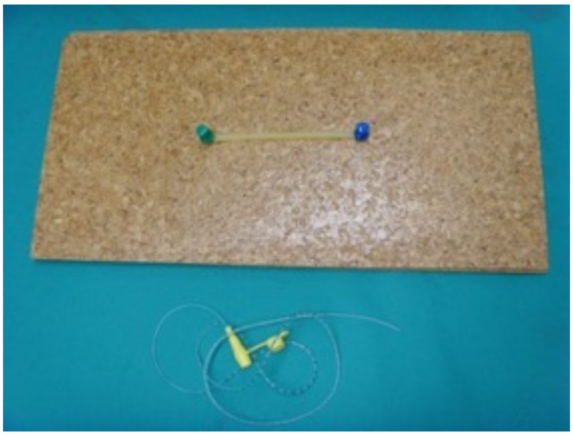
The pots are removed and on a board that has two crosses marked, labelled 1 and 2 a metal paper weight with a hook is placed over the left cross and a large metal paper clip placed in the middle.

The task is for the participant to grab hold of the paper clip with the graspers, hook the weight and then move it from 1 to 2 ten times, i.e. five return trips. Every time that the weight is lowered over the cross, (but not allowed to drag across the floor), the paper clip needs to be unhooked from the weight and hooked again. The paperclip should not be dropped by the graspers as this will account as an error. If the weight falls off the paperclip mid-air this will also count as an error.

The participant is explained that the time for speed and errors for accuracy will be measured as soon as the two instruments are in the box trainer and in view on the screen.

3.7.3 Laparoscopic Test 3: Bile duct cannulation

Figure 30: Bile Duct Cannulation



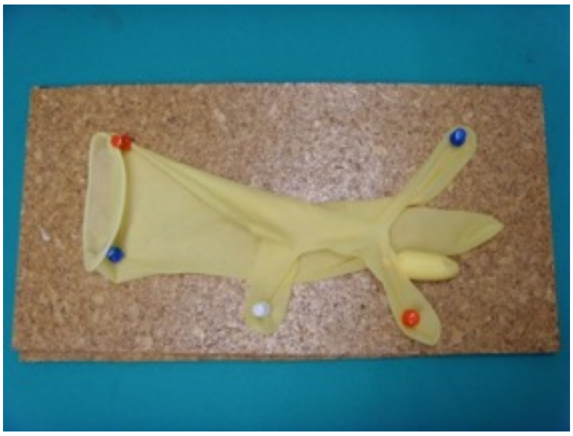
The cross labelled base is replaced by a cork board with Velcro on the bottom side that fixes itself to the floor of the box trainer. The cork board houses a 6.5 cm length of rubber catheter that is held down on both sides by tags and stretched slightly to make it 8 cm. In the middle of the catheter, on the upper side, a two third depth cut is made vertically down.

This is a more clinically relevant test than the previous two, where it is mimicking a clinical scenario. The objective is to pick up a clear and more rigid cannulation tube and insert it into the hole made in the fixed rubber catheter using a pair of graspers up to the first blue line that represents 5 cm.

The error will be if the force put into picking up the rubber catheter is so strong as to dislodge it from its cork base. The participant is told that both time and error will be measured and it will start as soon as both graspers are inside the box trainer and in view via the monitor.

3.7.4 Laparoscopic Test 4: Appendicectomy

Figure 31: Appendicectomy



Another cork base is placed, which carries a rubber glove size large. The large glove has a toothpaste filled tube inside the ring finger. At the base of this finger two rings have been drawn, half a centimetre apart, through which the participant, using a pair of scissors need to cut and free the toothpaste filled ring finger from the glove.

The glove is held down with tags, with the Index finger coming across the ring finger to cover the two lines that need cutting through. This requires the participant to use the graspers in the other hand to lift the Index finger to be able to complete the task. Both dominant and non-dominant hands are required to be used in this task more than any other. Errors in this task will be if any other part of the glove apart from in between the lines is cut, if the glove rips or comes off the board because too much force is applied pulling up.

The participant was able to choose whichever hand to place the scissors and the graspers. Time starts when both instruments are visible inside the box trainer on the screen.

3.8 Basic Surgical Skills Methods

26 participants were selected for the Basic Surgical Skills (BSS) tests. The selection consisted in picking at random five out of the top 10% Flying Aptitude Test scoring male participants and five females also from the top 10%. Another five males form the bottom 10% at the Flying Aptitude Test and five females from the bottom 10% were selected for this test. On the day of the test, seven participants pulled out at the last minute, which resulted in those slots to be filled by participants who had not yet taken the Flying Aptitude Test.

Open surgical skill tests were selected from the official Basic Surgical Skills Course. Each skill was shown three times on the standardised DVD used for the official BSS course where a stepwise approach detailing how to perform the task was shown.

Core skills were tested with a ‘Hand Tied Reef Knot’ (Figure 32) and an ‘Instrument Tie’ (interrupted suture): two scores were given for each test, one for the correct knot tied and the other for the correct tension applied. The same assessment scale used at the BSS course was used (Table 9).

Table 9: Basic Surgical Skills Assessment Sheet for Hand and Instrument Tie

	BSS Score	Comments	Signature of assessor
[redacted]			
Hand tied reef knot			
Correct knot tied			
Correct tension applied			
Instrument tie			
Correct knot tied			
Correct tension applied			

BSS Assessment of skill score

1. Inadequate, requires constant supervision
2. Adequate but needs further practice
3. Satisfactory

(MIN score 4, MAX score 12)

Figure 32: Hand Tie Station

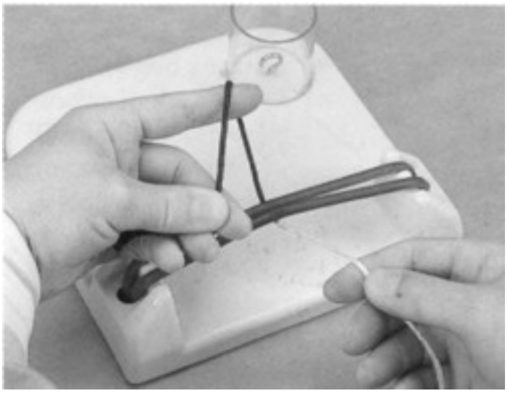
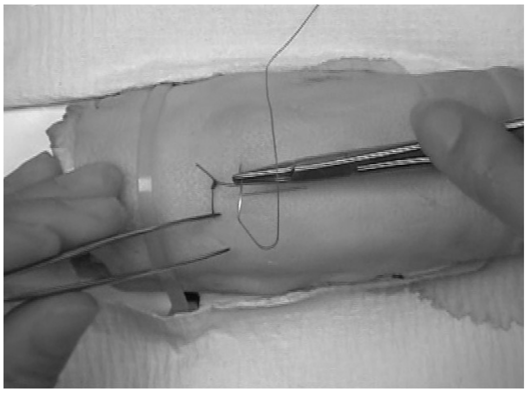


Figure 33: Suturing Station



There were two minor surgery tasks performed on an animal model (pig's trotters). First, the excision of a circular skin lesion using an elliptical incision (Table 10) and secondly the closure of the incision with interrupted sutures

Figure 33, Table 11). These tasks were more challenging than the knot tying and deserved a more comprehensive scoring system. The Global Rating Scale for Objective Structured Assessment of Technical Skills (OSATS Table 10,

Table 11) was used, which ranges from 1 to 5. It gives three scores for each assessment: Respect for tissue; time and motion; instrument handling.

The scoring was performed by two assessors using the above criteria and both the assessors were blind to the participants' aptitude tests and laparoscopic tests results.

Table 10: Excision Of Skin Lesion Global rating scale

	Rating				
Variable	1	2	3	4	5
Respect for tissue	Often used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately, with minimal damage
Time and motion	Many unnecessary moves		Efficient time and motion, but some unnecessary moves		Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments, although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness

MIN 3 MAX 15

Table 11: Suturing Global rating scale

	Rating				
Variable	1	2	3	4	5
Respect for tissue	Often used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately, with minimal damage
Time and motion	Many unnecessary moves		Efficient time and motion, but some unnecessary moves		Economy of movement and maximum efficiency
Instrument handling	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments, although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness

MIN 3 MAX 15

The total score of the four tasks, the 'Hand Tie', 'Instrument Tie', 'Suturing' and Excision of Skin Lesion ranged from a minimum of 9 to a maximum of 42.

3.9 Statistical Analysis

All data from the Flying Aptitude Test, Laparoscopic and Basic Surgical simulation tests as well as the questionnaire and demographic data was stored in Microsoft Office Excel 2011 (Microsoft USA). Data was then transferred to SPSS (IBM SPSS Statistics for Macintosh IBM Corp. Released 2011) for statistical analysis.

Shapiro-Wilk test and Q-Q plots were used to determine the normality of the data. before using Pearson correlation on the normally distributed Flying Aptitude Test and logged laparoscopic data and Spearman when correlating with the Basic Surgical Skills data as the latter was not normally distributed.

Independent T-Test was used when comparing data between the genders, under and post graduate participants and computer gamers and non-gamers.

Independent samples Kurskal-Wallis test was used when comparing the data between the four levels of gamer abilities.

Wilcoxon Signed Rank Test was used to compare the differences between the Self-Rating data and the results of the Flying Aptitude Test.

Chapter 4: RESULTS

This chapter described the results of the Flying Aptitude Test and the surgical laparoscopic and open simulated tests results. Gender, age, language, interest in a career in surgery and training grade differences as well as the use of computer games are described.

The undergraduates accounted for 86.6% (n=211) of the participants in this study out of n=243 and only data from the undergraduate participants were used for further analysis. By removing participants with prior surgical in vivo and simulation experience, data analysis from the 211 undergraduate participants will show a true cohort of complete military and surgical simulation novices.

Statistical analysis comparing different groups such as gender, age, training grade, English language, interest in a career in surgery or gastroenterology and the use of computer games were performed in the Flying Aptitude Tests, Laparoscopic and Basic Surgical Skill Tests.

4.1 Demographic Data

4.1.1 Training Grade

All participants had been through the British medical school system and either they were currently undergraduate medical students or practicing post-graduate doctors in the UK.

Most of the participants were medical students (n=211, 86.8%) with final (fifth) year students making the majority (n=128, 52.7%) (Table 12).

Foundation trainees (FY) have not entered specialty training yet, 14 out of the 15 core trainees (CT) year 1 were surgical and all four CT2 were surgical trainees. The specialty trainee (ST) Year 3 was a surgical academic and the other two ST4 and ST7 were gastroenterology medical senior trainees. Due to the variability of surgical experience amongst the postgraduate trainees the main statistical analysis of the data in this thesis were performed on the uniformly surgically novice group of undergraduate students (n=211).

Table 12: Training grade of participants

Training grade		Freq.	%	Freq.	%
Medical Student	Year 1	3	1.2	211	86.8
	Year 2	1	0.4		
	Year 3	44	18.1		
	Year 4	35	14.4		
	Year 5	128	52.7		
Foundation	FY1	7	2.9	10	4.1
	FY2	3	1.2		
Core Trainee	CT1	15	6.2	19	7.8
	CT2	4	1.6		
Specialty Trainee	ST3	1	0.4	3	1.2
	ST4	1	0.4		
	ST7	1	0.4		
	Total	243	100	243	100

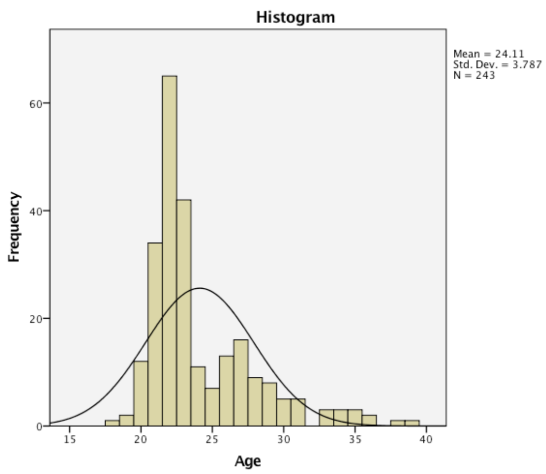
4.1.2 Gender

Out of the 243 participants 127 (52.3%) were male and 116 (47.7%) were female. The age range for the 243 participants was from 18 to 39 with the mean at 24.15 and median of 23 years old. Both the youngest and oldest participant were undergraduate students, the oldest being a graduate entry medical student (Table 13). The age of the participants was not normally distributed in both undergraduate and postgraduate (Figure 34).

Table 13: Age distribution of all participants, UG & PG

Age distribution		ALL		UG		PG	
	n	243		211		32	
		♀116	♂127	♀107	♂104	♀9	♂23
Mean		24.11		23.52		28.16	
95% CI for mean	Upper	23.63		23.06		26.72	
	Lower	24.59		23.97		29.59	
Median		23		22		28	
Mode		22		22		28	
SD		3.79		3.36		3.98	
Variance		14.339		11.31		15.81	
Range		21		21		15	
IQR		4		2		6	
Min		18		18		23	
Max		39		39		38	
Percentiles	25	25		22		25	
	50	23		22		28	
	75	26		24		30.75	

Figure 34: All Participants age distribution



4.2 Flying Aptitude Test Results

4.2.1 Demographic Data

230 participated in the Flying Aptitude Test. The Flying Aptitude Test Index score (a percentage) having derived from combining the five aptitude tests which are given individual weighting to match the chosen task of laparoscopic cholecystectomy: Psychomotor 40%, Verbal Reasoning 20%, Spatial Reasoning 20%, Attentional Capability 10% and Short Term Memory 10%.

Table 14: Flying Aptitude Test scores

	Index		Psycho-motor		Verbal Reasoning		Attentional capability		Spatial Reasoning		Short Term Memory	
	ALL	UG	ALL	UG	ALL	UG	ALL	UG	ALL	UG	ALL	UG
N	230	199	230	199	230	199	230	199	230	199	230	199
Mean	51.78	51.64	4.07	3.96	5.75	5.86	4.04	4.08	4.99	4.98	4.73	4.79
95% CI	49.88	49.64	3.82	3.7	5.51	5.61	3.75	3.77	4.75	4.73	4.48	4.52
of mean	53.68	53.63	4.32	4.23	6	6.12	4.34	4.39	5.22	5.23	4.99	5.06
Median	50	50	4	4	6	6	4	4	5	5	5	5
Mode	49	40	4	4	5	5	1	5	6	6	5	5
SD	14.63	14.27	1.93	1.88	1.87	1.81	2.26	2.22	1.81	1.81	1.94	1.93
Variance	214.13	203.75	3.74	3.55	3.5	3.29	5.09	4.92	3.28	3.28	3.78	3.73
Range	80	80	8	8	8	8	8	8	8	8	8	8
IQR	19	18	2	0.61	2	-0.16	3	0.31	2	-0.11	3	-0.02
Skewness	0.3	0.27	0.61	0.13	-0.17	-0.42	0.34	-0.58	-0.11	-0.49	-0.14	-0.75
Kurdosis	-0.44	-0.08	0.84	8.00	-0.5	8.00	-0.68	8.00	-0.57	8.00	-0.79	8.00
Min	16	16	1	1	1	1	1	1	1	1	1	1
Max	96	96	9	9	9	9	9	9	9	9	9	9
Perc 25%	42	25	3	42	5	3	2	4	4	5	3	3
50%	50	50	4	4	6	6	4	4	5	5	5	5
75%	61	60	5	5	7	7	5	5	6	6	6	6

UG: undergraduate participants only

The median score (STANINE scoring system described in Chapter 3.5) for all five aptitudes the median lied between 4 and 6 making our medical candidates results comparable to the RAF candidates. The median value for the

Psychomotor Aptitude is 4, which is the heaviest weighted aptitude in our chosen task (Table 14).

Only the Flying Aptitude Test Index (final percentage score with weighted individual aptitudes) showed a normal distribution on the Shapiro-Wilk test. The rest of the five aptitudes were not normally distributed (Table 15).

The mean for the Flying Aptitude Test Index score for all participants (n=230) was 51.78 with a minimum score of 16 and maximum of 96 (Table 14) and it showed a normal distribution visible on histogram (Figure 35) and a negative Shapiro-Wilk normality Test (Table 15). The range of all five aptitudes included the top 4% (scoring 9) and also to the bottom 4% (scoring 1) when compared to the rest of the RAF candidates.

Undergraduate participants accounting for 86.5% (n=199) showed the same normal distribution curve and normal Q-Q Plot (Figure 36).

Table 15: Shapiro-Wilk Test for Normal distribution

Tests of Normality	All Participants N= 230		Undergraduates N= 199		Postgraduates N=31	
	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.
Index	0.988	0.052	0.989	0.142	0.96	0.3
Psychomotor	0.941	<0.001	0.943	<0.001	0.94	0.085
Verbal Reasoning	0.962	<0.001	0.961	<0.001	0.957	0.242
Spatial reasoning	0.961	<0.001	0.962	<0.001	0.94	0.081
Attentional capability	0.934	<0.001	0.934	<0.001	0.888	<0.014
Short Term Memory	0.962	<0.001	0.963	<0.001	0.954	0.203

Figure 35: Frequency Distribution Flying Aptitude Test Index (%) Results for all Participants

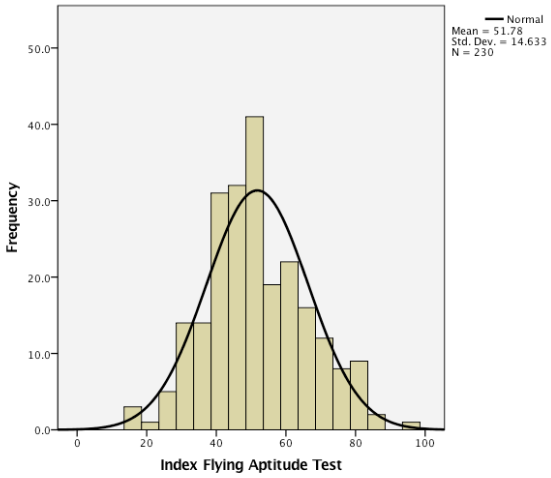
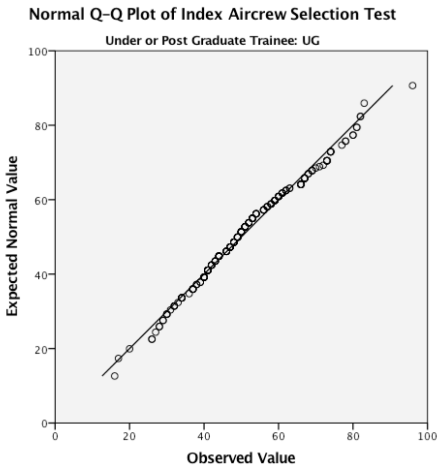


Figure 36: Normal Q-Q Plot of Index score (UG)



Independent t-tests mean comparison were performed for the Flying Aptitude Test Scores between the undergraduate (UG n=199) and postgraduate (PG n=31) participants. The only statistically significant aptitude was the Verbal Reasoning where UG scored higher than the PG (UG 5.86 vs. 5.03 PG p=0.021). The final Index score (UG 51.64 vs. 52.71 PG p=0.741) did not show a statistical significant difference between the two groups (Figure 37, Figure 38, Figure 39).

Figure 37: UG Flying Aptitude Test Index Scores

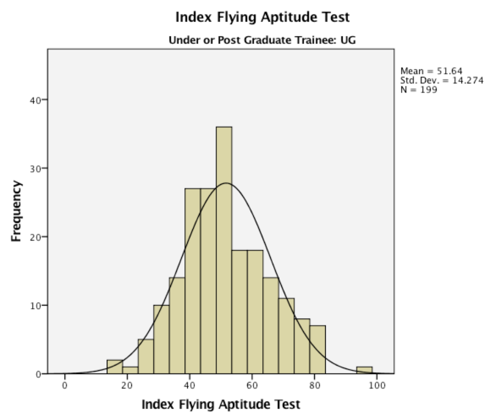


Figure 38: PG Flying Aptitude Test Index Scores

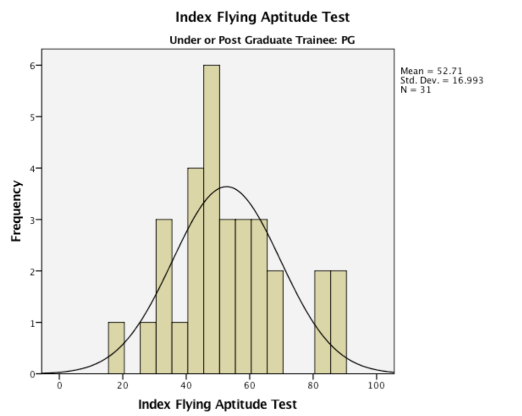
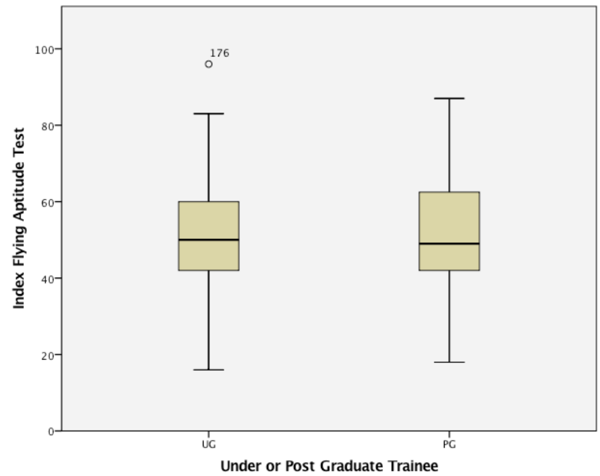


Figure 39: Boxplot Flying Aptitude Test Index score UG & PG



4.2.2 Correlation between the five Flying Aptitude Tests

Pearson Correlation between the undergraduates' five aircrew selection domains results (Table 16) have shown a statistically significant positive correlation between all of the five aptitude domains. The highest correlation was between Psychomotor and Attentional Capability at r=0.499 (p<001). Verbal Reasoning had the lowest correlation value but still significant.

Table 16: Pearson Correlation between the Aptitude Tests

n=199		Psychomotor	Verbal Reasoning	Spatial reasoning	Attentional capability	Short Term Memory
Psychomotor	Pearson r	1	.164*	.282**	.499**	.280**
	Sig. (2-tailed)		0.021	<0.001	<0.001	<0.001
Verbal Reasoning	Pearson r	.164*	1	.144*	.147*	.329**
	Sig. (2-tailed)	0.021		0.043	0.038	<0.001
Spatial reasoning	Pearson r	.282**	.144*	1	.410**	.261**
	Sig. (2-tailed)	<0.001	0.043		<0.0010	<0.001
Attentional capability	Pearson r	.499**	.147*	.410**	1	.298**
	Sig. (2-tailed)	<0.001	0.038	<0.001		<0.001
Short Term Memory	Pearson r	.280**	.329**	.261**	.298**	1
	Sig. (2-tailed)	<0.001	<0.001	<0.001	<0.001	

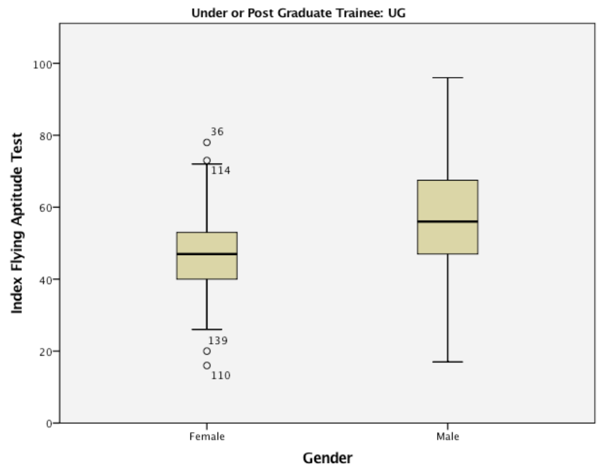
*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

4.2.3 Gender Differences in the Flying Aptitude Tests

Undergraduate participants' Flying Aptitude Test Index score box plot shows that a difference in the mean between the genders but the range being wider in males (Figure 40).

Figure 40: Undergraduate Flying Aptitude Test Index score ♀:♂



The Psychomotor aptitude showed the greatest statistical difference in the distribution of the scores between the genders (Figure 41). The psychomotor component was the heaviest weighted aptitude (40%) in the Index score which influenced it with a 10% difference in distribution between the genders (mean female 46.51 vs. 57.14 male p<0.001) (Table 17).

Figure 41: UG Psychomotor Domain ♀:♂

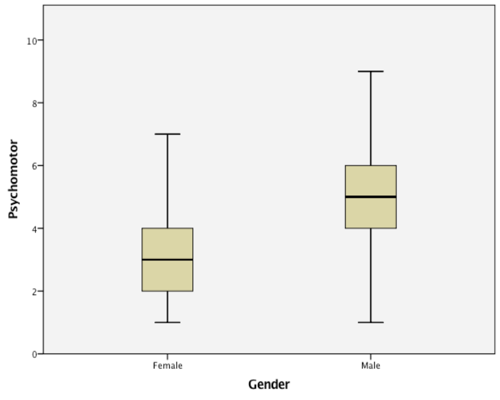


Figure 42: Attentional capabilities ♀:♂

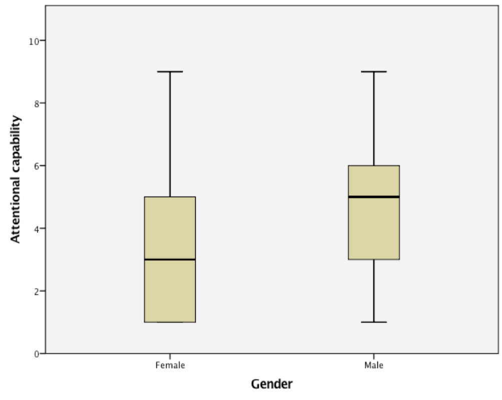


Table 17: Undergraduate Independent t-test of Means in Flying Aptitude Tests between genders

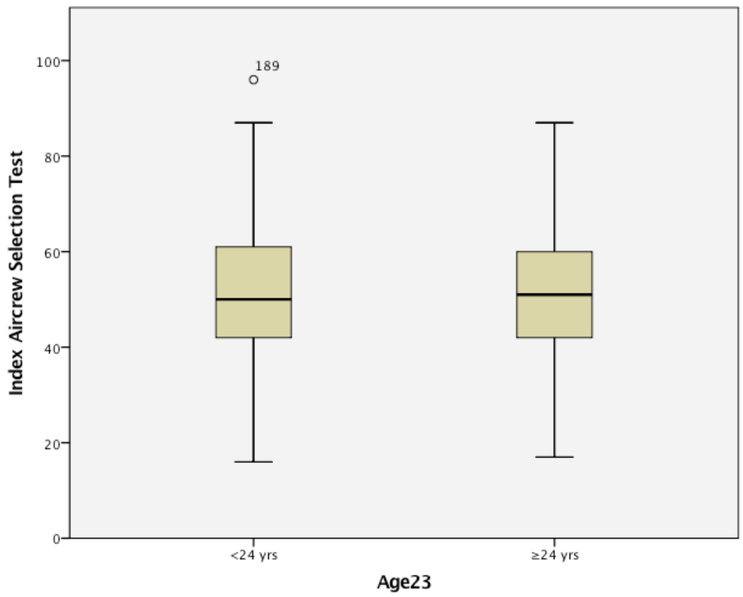
Total n=199	Gender	n	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Index	Female	103	46.51	11.835	1.166	
	Male	96	57.14	14.672	1.497	<0.001
Psychomotor	Female	103	2.93	1.285	0.127	
	Male	96	5.07	1.796	0.183	<0.001
Verbal Reasoning	Female	103	5.99	1.917	0.189	
	Male	96	5.73	1.695	0.173	0.311
Spatial Reasoning	Female	103	4.95	1.688	0.166	
	Male	96	5.01	1.944	0.198	0.819
Attentional Capability	Female	103	3.44	2.099	0.207	
	Male	96	4.76	2.146	0.219	<0.001
Short Term Memory	Female	103	4.79	1.913	0.188	
	Male	96	4.79	1.962	0.2	0.985

Independent t-test of means between female and male participants was conducted on all Flying Aptitude Tests (Table 17). Male participants had a higher mean score on the Psychomotor (♀ 2.93 vs 5.07 ♂; p<0.001) and Attentional Capability ((♀ 3.44 vs 4.76 ♂; p<0.001).

4.2.4 Age Differences

There was no statistical difference between the mean results of the five individual Flying Aptitude and the final Index score when comparing the under 24 year olds (n=145) and the 24 and over year old participants (n=85) (Figure 43). The age of 23 was chosen to compare the Graduate Entry Medicine (GEM) Students to the straight from school into medical school cohort. The paper published by Fielding et al [3] showed that Graduate Entry Students were consistently slower in completing all of the laparoscopic simulated tasks than the undergraduates straight from school.

Figure 43: Flying Aptitude Test Index Score difference between under and over 24 years



4.3 Laparoscopic Simulation Tests

4.3.1 Demographic data

Data from 160 undergraduates were collected for the laparoscopic simulator tests in the box trainer. The mean value for the Total Laparoscopic Time was 737.99 seconds with a tenfold range (min 259sec to max 2290sec; range 2031sec) (Table 18). The data distribution on a histogram showed a positive Skewness (Figure 44) and this was also true for the distribution of the four individual tests.

Table 18: Laparoscopic Simulator Tests Data Statistics

N=160		Total lap time	Bean drop	Block move	Bile duct cannulation	Appendicec-tomy
Mean		737.99	170.61	228.59	185.2	153.6
95% CI mean	Lower	689.12	159.32	206.38	156.58	137.86
	Upper	786.87	181.89	250.79	213.82	169.34
Median		689	151	184.5	114.5	123.5
Mode		396[b]	104[b]	148	54[b]	50[b]
SD		313.018	72.262	142.227	183.287	100.795
Variance		97980.321	5221.813	20228.609	33594.262	10159.726
Range		2031	477	1085	904	536
IQR		378	86	158	165	131
Min		259	73	57	14	14
Max		2290	550	1142	918	550
%	25	525.5	118.5	135.5	70	78
	50	689	151	184.5	114.5	123.5
	75	903.25	204	293.75	234.75	208.5

b. Multiple modes exist. The smallest value is shown

The normal Q-Q plot of the Total Laparoscopic Times and the Shapiro-Wilk Test for normality demonstrate that the data is not normally distributed. To be able to correlate this data with the normally distributed Flying Aptitude Tests results the Laparoscopic data was transformed into Log Data.

Figure 44: UG Total Laparoscopic Simulation Test Times

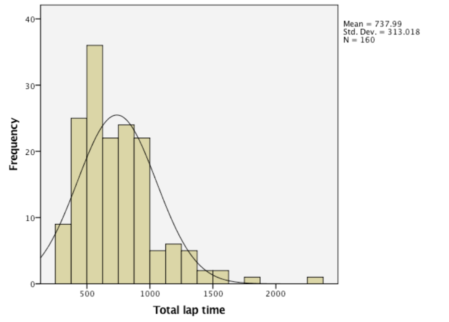


Figure 45: UG Log Transformed Total Lap Sim Test Times

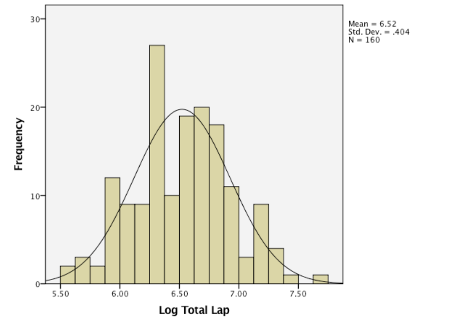


Figure 46: Normal Q-Q Plot Total Lap Time

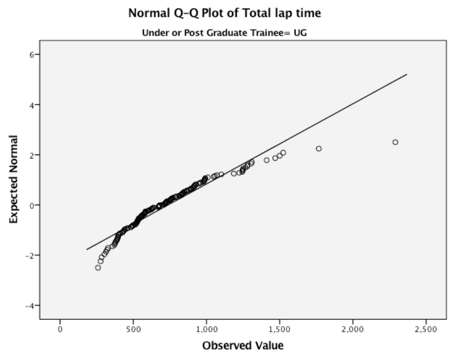
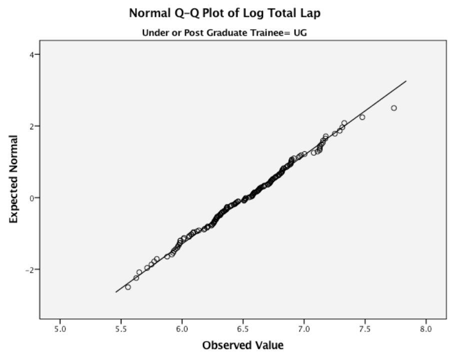


Figure 47: Log Transformed Q-Q Plot Total Lap Time



Log Transform of the Laparoscopic Total Time Results has produced a normal distribution histogram (Figure 45) and Q-Q Plot (Figure 47). Table 19 shows the Shapiro-Wilk Test of Normality for the original Laparoscopic Simulated Tests results and the normalised Log Transformed Data.

Table 19: Test of Normality Shapiro-Wilk Laparoscopic Simulator Test. Normal & Log Data

	df	Statistic	Statistic	df	Statistic
Total lap time	0.912	<0.001	Log VR Lap Total Time	0.984	0.982
Bean drop	0.861	<0.001	Log Bean Drop	0.914	0.345
Block move	0.806	<0.001	Log Block Move	0.939	0.568
Bile duct cannulation	0.764	<0.001	Log Appendicectomy	0.942	0.601
Appendicectomy	0.897	<0.001	Log BDC	0.949	0.676

Both Bean Drop and Block move have a significant correlation with all the laparoscopic simulator tests but the two more clinically oriented tests, bile duct cannulation and appendicectomy had the lowest correlation at 0.08 and did not show statistical significance (p=0.313) (Table 20).

Table 20: Pearson Correlation between the four laparoscopic simulated tests and the total score

Pearson Correlation N=160		Log Total Lap	Log Bean Drop	Log Block Move	Log Appendi-cectomy	Log BDC
Log Total Lap	r	1	.627**	.631**	.479**	.716**
	Sig.		<0.001	<0.001	<0.001	<0.001
Log Bean Drop	r	.627**	1	.339**	.282**	.319**
	Sig.	<0.001		<0.001	<0.001	<0.001
Log Block Move	r	.631**	.339**	1	.160*	.166*
	Sig.	<0.001	<0.001		0.044	0.036
Log Appendicectomy	r	.479**	.282**	.160*	1	0.08
	Sig.	<0.001	<0.001	0.044		0.313
Log BDC	r	.716**	.319**	.166*	0.08	1
	Sig.	<0.001	<0.001	0.036	0.313	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

The strongest correlation was between the two non-clinically oriented tests: Bean Drop and Block Move (0.339 p<<0.001) and the Bean Drop had the highest overall correlation scores with the rest of the three tests. The total laparoscopic time had significant correlation with all four tests with the highest being Bile Duct Cannulation at 0.716 (p<0.001). The Block Move test had a mean of 229 seconds. With the range being the widest of all the tests at 1085 seconds (57-1142 sec). The Bile Duct Cannulation and Appendicectomy tasks had the fastest times recorded at 14 seconds each. The Bile Duct Cannulation test had a mean time of 185 seconds (14-918; range=904 sec). The Appendicectomy test had a mean of 154 seconds (14-550; range=536).

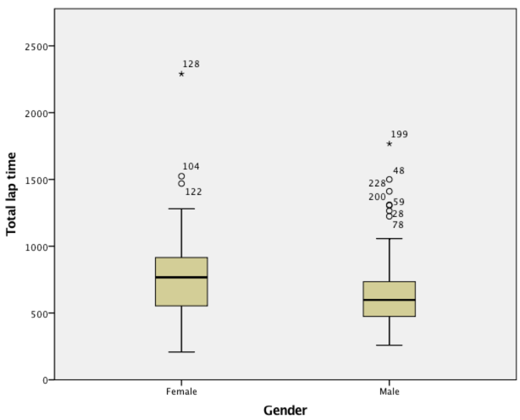
4.3.2 Gender Differences: Laparoscopic Simulation Tests

Female participants (n=81) generally completed the tasks in a longer time than male participants (n=79) but only the Bean Drop test out of the four showed a statistical significance (Female 188 vs. 153 seconds Male p=<0.012). This had an influence on the total lap time which also showed a difference where Females had a mean of 800 seconds vs. 674 seconds for Males p=0.01 (Table 21, Figure 48)

Table 21: Independent T-test for comparing the Means of laparoscopic simulation tests between the genders.

	Gender	N	Mean	SD	Std. Error Mean	Sig. (2-tailed)
Total Lap Sim Time	Female	81	800.41	312.221	34.691	
	Male	79	674	302.58	34.043	0.010
Bean drop	Female	81	187.78	81.319	9.035	
	Male	79	153	56.939	6.406	<0.012
Block move	Female	81	248.73	153.436	17.048	
	Male	79	207.94	127.42	14.336	0.069
Bile duct cannulation	Female	81	208.53	188.256	20.917	
	Male	79	161.28	176.031	19.805	0.103
Appendicectomy	Female	81	155.37	94.017	10.446	
	Male	79	151.78	107.879	12.137	0.823

Figure 48: Total Lap Time Mean Score ♀:♂



4.3.2 Flying Aptitude Test & Laparoscopic Tests

153 undergraduate participants (female 78; male 75) underwent both the Flying Aptitude Test and the four Laparoscopic Surgical Simulated (Lap Sim) tests. The log values were calculated to normalise the data of the four Lap Sim tests and the Total Lap Sim time. These have been correlated (Pearson's) with the log values of the five Flying Aptitude Tests and the final Index score (Table 22).

Table 22: Pearson Correlation between Flying Aptitude Tests and Log Laparoscopic Simulation Tests.

		Log Bean Drop	Log Block Move	Log BDC	Log Appendicectomy	Log Total Lap
Psychomotor	Pearson Correlation	-.270**	-.186*	-.141	-.187*	-.300**
	Sig. (2-tailed)	.001	.021	.083	.021	.000
	N	153	153	153	153	153
Verbal Reasoning	Pearson Correlation	-.054	.050	-.133	-.009	-.057
	Sig. (2-tailed)	.509	.543	.100	.912	.483
	N	153	153	153	153	153
Spatial reasoning	Pearson Correlation	-.104	-.145	-.005	-.076	-.113
	Sig. (2-tailed)	.202	.074	.949	.352	.166
	N	153	153	153	153	153
Attentional capability	Pearson Correlation	-.239**	-.221**	-.188*	.011	-.235**
	Sig. (2-tailed)	.003	.006	.020	.897	.003
	N	153	153	153	153	153
Short Term Memory	Pearson Correlation	-.035	-.001	-.080	-.006	-.073
	Sig. (2-tailed)	.665	.986	.327	.938	.370
	N	153	153	153	153	153
Index Aircrew Selection Test	Pearson Correlation	-.244**	-.169*	-.171*	-.132	-.275**
	Sig. (2-tailed)	.002	.036	.034	.104	.001
	N	153	153	153	153	153

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

The Psychomotor domain showed a significant correlation with all the four Lap Sim tests and therefore with the Total Lap Sim time. The highest significant correlation value being between the Psychomotor domain and the Bean Drop Test (r=-0.270 p=<0.011). The correlation between the Flying

Aptitude Test Index score and the Total Lap Sim was r=-0.274 p=<0.011 (Figure 49).

Figure 49: Total Lap Sim & Flying Aptitude Test Index Score Correlation

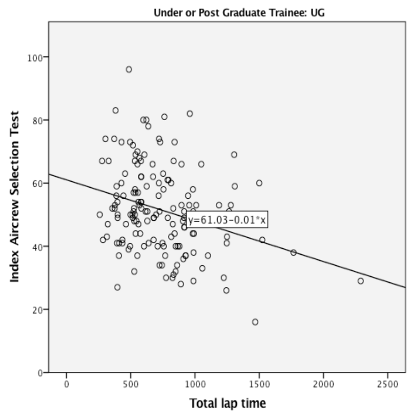
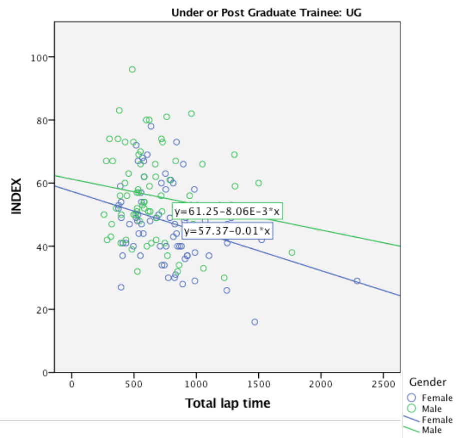


Figure 50: Gender specific Total Lap Sim & Flying Aptitude Test Index Correlation



Gender subgroup data was analysed and the female participants did show a statistically significant correlation between the Flying Aptitude Test Index score and Total Lap time (r=-0.288 p=0.011) as well as for the Psychomotor (r =-0.226; p=0.047) and Attentional Capability (r=-0.259; p=0.022) domains (Table 24). However, there was no statistically significant correlation between the male participants' Lap Sim tests and the Flying Aptitude Test results (Flying Aptitude Test Index score and Log Total Lap time Pearson correlation (r=-0155 p=0.183)) (Table 25).

No statistical significance was found in the results between the genders when the two correlation coefficients in the two independent samples were analysed using the Fisher transformation test (Table 23).

Table 23: Fisher Transformation of Correlation Coefficients between genders (F.A.T & Lap Sim)

Gender	Correlation Coefficient	Fisher r to z transformation	Significance	
♀ n = 78	-0.288	z = -0.85	0.1977	1 tailed
♂ n = 75	-0.155		0.3953	2 tailed

Table 24: Pearson Correlation UG Female Flying Aptitude Test & Lap Sim

		Log Bean Drop	Log Block Move	Log BDC	Log Appendicectomy	Log Total Lap
Psychomotor	Pearson Correlation	-.177	-.142	-.011	-.225*	-.226*
	Sig. (2-tailed)	.121	.215	.925	.048	.047
	N	78	78	78	78	78
Verbal Reasoning	Pearson Correlation	-.105	.047	-.161	-.014	-.093
	Sig. (2-tailed)	.360	.685	.159	.906	.420
	N	78	78	78	78	78
Spatial reasoning	Pearson Correlation	-.122	-.163	-.129	-.053	-.204
	Sig. (2-tailed)	.288	.153	.260	.644	.074
	N	78	78	78	78	78
Attentional capability	Pearson Correlation	-.102	-.217	-.207	-.057	-.259*
	Sig. (2-tailed)	.374	.056	.068	.623	.022
	N	78	78	78	78	78
Short Term Memory	Pearson Correlation	-.051	-.042	-.152	-.083	-.159
	Sig. (2-tailed)	.659	.714	.185	.472	.163
	N	78	78	78	78	78
Index Aircrew Selection Test	Pearson Correlation	-.191	-.157	-.170	-.157	-.288*
	Sig. (2-tailed)	.094	.169	.136	.169	.011

Table 25: Pearson Correlation UG Male Flying Aptitude Test & Lap Sim

		Log Bean Drop	Log Block Move	Log BDC	Log Appendicectomy	Log Total Lap
Psychomotor	Pearson Correlation	-.130	-.102	-.064	-.177	-.190
	Sig. (2-tailed)	.265	.385	.583	.128	.103
	N	75	75	75	75	75
Verbal Reasoning	Pearson Correlation	-.070	.014	-.161	-.014	-.087
	Sig. (2-tailed)	.548	.905	.167	.903	.459
	N	75	75	75	75	75
Spatial reasoning	Pearson Correlation	-.141	-.157	.085	-.100	-.079
	Sig. (2-tailed)	.228	.178	.469	.392	.500
	N	75	75	75	75	75
Attentional capability	Pearson Correlation	-.260*	-.151	-.062	.097	-.097
	Sig. (2-tailed)	.024	.196	.600	.406	.406
	N	75	75	75	75	75
Short Term Memory	Pearson Correlation	-.052	.022	-.029	.055	-.023
	Sig. (2-tailed)	.659	.852	.808	.640	.842
	N	75	75	75	75	75
Index Aircrew Selection Test	Pearson Correlation	-.156	-.099	-.067	-.102	-.155
	Sig. (2-tailed)	.181	.396	.570	.386	.183
	N	75	75	75	75	75

4.4 Basic Surgical Skills Tests Results

4.4.1 Demographic Data

22 undergraduate participants took part in the Basis Surgical Skills (BSS) (9 Female, 13 Male). The mean age was 23 years (20-28 years). The mean total BSS score was 70.17% and median 81.25%. The range is quite wide from 15.63% to 93.75%.

Table 26: Frequency distribution BSS data

		Mean	Median	Mode	SD	Variance	Range	Min	Max
Hand Tie (1-3)	Correct Tie	2.5	3	3	0.74	0.55	2	1	3
	Tension	2.32	2	3	0.72	0.51	2	1	3
Instrument Tie (1-3)	Correct Tie	2.86	3	3	0.35	0.12	1	2	3
	Tension	2.75	3	3	0.43	0.19	1	2	3
Tie score BSS system (4-12)		10.43	11	12	1.82	3.29	6	6	12
Suturing (1-5)	Tissue Handling	3.45	4	4	1.21	1.47	4	1	5
	Time & Motion	3.45	3.5	3	1.05	1.09	4	1	5
	Instrument handling	3.32	3	3	1.16	1.35	4	1	5
Lesion Excision (1-5)	Tissue Handling	3.91	4	5	1.04	1.09	4	1	5
	Time & Motion	4.02	4.25	5	1.02	1.04	3	2	5
	Instrument handling	3.86	4.5	5	1.39	1.93	4	1	5
suturing & lesion score OSATS (6-30)		22.02	24.5	26	6.05	36.56	21	8	29
Total BSS score (10-42)		32.45	36	36	7.46	55.71	25	15	40
Total BSS % score		70.17	81.25	81.25	23.33	544.06	78.12	15.63	93.75

The frequency histogram shows a negative skewed graph on both the raw scores (Figure 51) and the percentage scores (Figure 52). The attempt to normalise the data failed after Log transform was run (supported by histogram distribution and Shapiro-Wilks test for normalisation). Therefore, the original data for the BSS was used and Spearman correlation performed instead.

Figure 51: BSS freq. total % scores

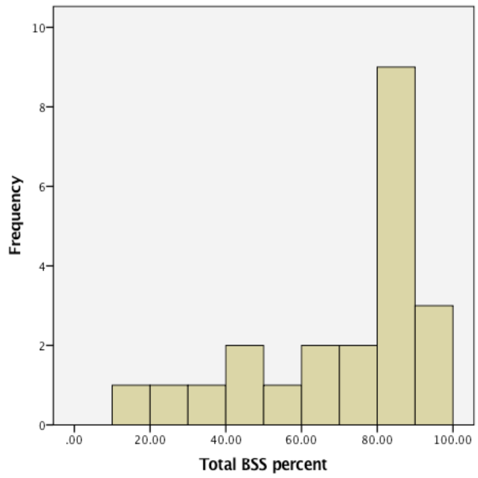
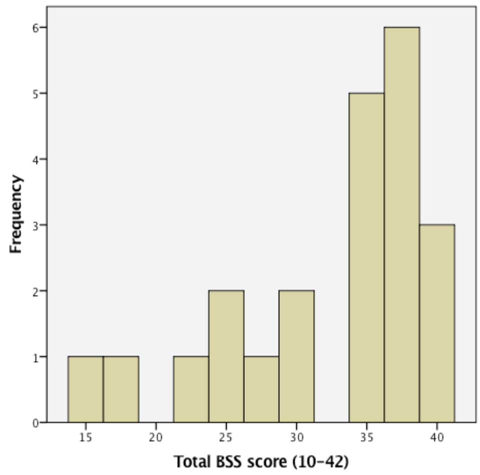


Figure 52: BSS freq. total raw scores



Correlation between the ten BSS Tests: There was a positive correlation between the total score for the three point BSS 'Hand Tie' and 'Instrument Tie' scores and the five point OSATS results (Spearman Correlation r=0.518, p=0.01).

The two 'Hand Tie' scores for 'correctly tying' and 'tension' correlated well (r=0.796; p<0.01) but 'Instrument Tie' did not. For the three subdivision of scores for the 'Suturing' and 'Lesion Excision' excision and amongst all other scores the correlation was high (r= 0.5~0.8) and statistically significant (Table 27).

Table 26

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 27: Basic Surgical Skills Spearman Correlation between tests (n=22)

Spearman Correlation			Hand Tie		Instrument Tie		Tissue	Suturing			Lesion Excision			OSATS	Total
UG n=22			Correct	Tension	Correct	Tension	Total BSS	Tissue	T&M	Inst	Tissue	T&M	Inst	Total	
Hand Tie	Correct tie	r	1	.796**	0.37	0.27	.827**	.454*	.557**	0.41	0.26	0.30	0.37	0.40	.542**
		Sig.	.	<0.01	0.09	0.22	<0.01	0.03	0.01	0.06	0.25	0.17	0.09	0.07	0.01
	Tension	r	.796**	1	0.30	.493*	.932**	.618**	.656**	.675**	.442*	.450*	.539**	.609**	.735**
		Sig.	<0.01	.	0.18	0.02	<0.01	<0.01	<0.01	<0.01	0.04	0.04	0.01	<0.01	<0.01
Instrument Tie	Correct tie	r	0.37	0.30	1	0.30	0.42	0.26	0.28	0.20	0.19	0.20	0.23	0.18	0.26
		Sig.	0.09	0.18	.	0.18	0.06	0.24	0.21	0.36	0.39	0.38	0.30	0.43	0.24
	Tension	r	0.27	.493*	0.30	1	.594**	.538**	.485*	.506*	.542**	.508*	.611**	.516*	.507*
		Sig.	0.22	0.02	0.18	.	<0.01	0.01	0.02	0.02	0.01	0.02	<0.01	0.01	0.02
Tie total (BSS)		r	.827**	.932**	0.42	.594**	1	.597**	.635**	.598**	0.35	0.42	.502*	.518*	.667**
		Sig.	<0.01	<0.01	0.06	<0.01	.	<0.01	<0.01	<0.01	0.11	0.06	0.02	0.01	<0.01
Suturing	Tissue handling	r	.454*	.618**	0.26	.538**	.597**	1	.768**	.776**	.498*	.710**	.775**	.869**	.857**
		Sig.	0.03	<0.01	0.24	0.01	<0.01	.	<0.01	<0.01	0.02	<0.01	<0.01	<0.01	<0.01
	Time & Motion	r	.557**	.656**	0.28	.485*	.635**	.768**	1	.864**	0.39	.581**	.640**	.835**	.882**
		Sig.	0.01	<0.01	0.21	0.02	<0.01	<0.01	.	<0.01	0.07	0.01	<0.01	<0.01	<0.01
	Instrument handling	r	0.41	.675**	0.20	.506*	.598**	.776**	.864**	1	.548**	.647**	.650**	.893**	.904**
		Sig.	0.06	<0.01	0.36	0.02	<0.01	<0.01	<0.01	.	0.01	<0.01	<0.01	<0.01	<0.01
Lesion Excision	Tissue handling	r	0.26	.442*	0.19	.542**	0.35	.498*	0.39	.548**	1	.764**	.747**	.725**	.679**
		Sig.	0.25	0.04	0.39	0.01	0.11	0.02	0.07	0.01	.	<0.01	<0.01	<0.01	<0.01
	Time & Motion	r	0.30	.450*	0.20	.508*	0.42	.710**	.581**	.647**	.764**	1	.925**	.838**	.786**
		Sig.	0.17	0.04	0.38	0.02	0.06	<0.01	0.01	<0.01	<0.01	.	<0.01	<0.01	<0.01
	Instrument handling	r	0.37	.539**	0.23	.611**	.502*	.775**	.640**	.650**	.747**	.925**	1	.840**	.794**
		Sig.	0.09	0.01	0.30	<0.01	0.02	<0.01	<0.01	<0.01	<0.01	<0.01	.	<0.01	<0.01
Suturing & Excision Total (OSATS)		r	0.40	.609**	0.18	.516*	.518*	.869**	.835**	.893**	.725**	.838**	.840**	1	.972**
		Sig.	0.07	<0.01	0.43	0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	.	<0.01
Total BSS score (10-42)		r	.542**	.735**	0.26	.507*	.667**	.857**	.882**	.904**	.679**	.786**	.794**	.972**	1
		Sig.	0.01	<0.01	0.24	0.02	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	.

4.4.2 Gender differences in Basic Surgical Skills Tests

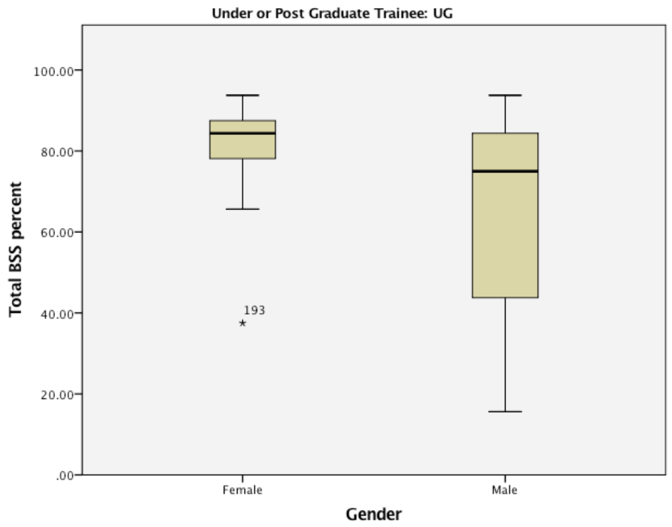
The Independent T-Test comparison of means between the genders in the total Basic Surgical Skills score showed the female undergraduate (n=9) participants scored 14% higher than male participants (n=13) (78.47% vs. 64.42%); p=0.170) (Table 28, Appendix 12). Every single mean value of the ten scoring parameters in the BSS showed that the female participants score higher than the males but the independent sample T-Test did not show statistical significance.

Table 28: Independent T-Test between male and female for BSS tests

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Hand tie Correct	Female	9	2.56	.726	.242
	Male	13	2.46	.776	.215
Hand tie tension	Female	9	2.44	.726	.242
	Male	13	2.23	.725	.201
Instrument tie Correct	Female	9	2.78	.441	.147
	Male	13	2.92	.277	.077
Instrument tie tension	Female	9	2.83	.354	.118
	Male	13	2.69	.480	.133
Suturing tissue handling	Female	9	4.00	.707	.236
	Male	13	3.08	1.367	.379
Suturing Time & Motion	Female	9	3.72	.905	.302
	Male	13	3.27	1.129	.313
Suturing Instrument handling	Female	9	3.72	1.034	.345
	Male	13	3.04	1.198	.332
Lesion excision Tissue handling	Female	9	4.28	.972	.324
	Male	13	3.65	1.049	.291
Lesion excision Time & Motion	Female	9	4.39	.993	.331
	Male	13	3.77	.992	.275
Lesion Instrument handling	Female	9	4.39	.993	.331
	Male	13	3.50	1.541	.427
Tie score using BSS system (4-12)	Female	9	10.61	1.728	.576
	Male	13	10.31	1.932	.536
suturing & lesion score using OASTS system (6-30)	Female	9	24.50	4.514	1.505
	Male	13	20.31	6.527	1.810
Total BSS score (10-42)	Female	9	35.11	5.578	1.859
	Male	13	30.62	8.234	2.284
Total BSS percent	Female	9	78.4744	17.43059	5.81020
	Male	13	64.4238	25.73043	7.13634

Data on the box plot graph shows the male participants had a much wider range of results than the female participants (Figure 53). The standard deviation for the male participants was 25.73 compared to 17.43 in females. The lowest scoring two participants were male.

Figure 53: Gender differences BSS results



4.4.3 Flying Aptitude Test and Basic Surgical Skills

20 undergraduate participants had undertaken both the Aircrew Selection Test (AST) and the Basic Surgical Skills test of whom 8 were female and 12 male. Table 27 shows the Spearman correlation between the ten BSS and its total scores and the five Flying Aptitude test domains and the Index scores. The Flying Aptitude Test Index scores and the total BSS scores showed a positive correlation of r = 0.376 (Figure 54), but the p value was outside of significance p=0.103. The reason may be due to the non-existent correlation between the 3

point scale Tying Tests ('Hand Tie' and 'Instrument Tie') and the Flying Aptitude Test result (r=0.056; p=0.814) (Figure). The correlation between the total scores of the both five-point OSATS scale tests (the 'Suturing' and 'Lesion Excision' Tests) showed a positive correlation with the Flying Aptitude Test Index scores of r=0.464; p=0.04 (

Figure 56).

Figure 54: BSS and Flying Aptitude Test Index scores scatter plot

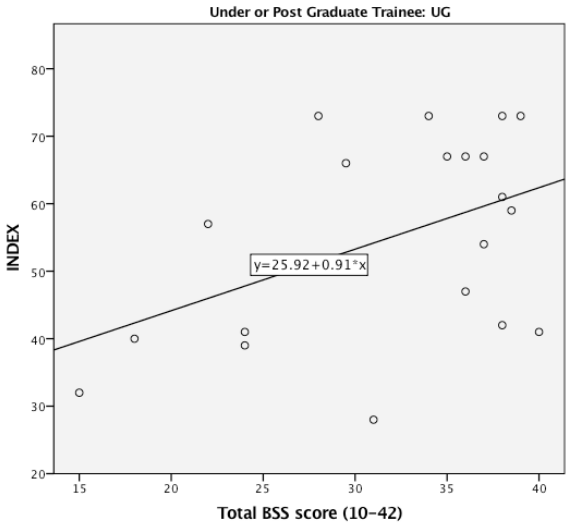


Figure 55: Figure: Tying & Index

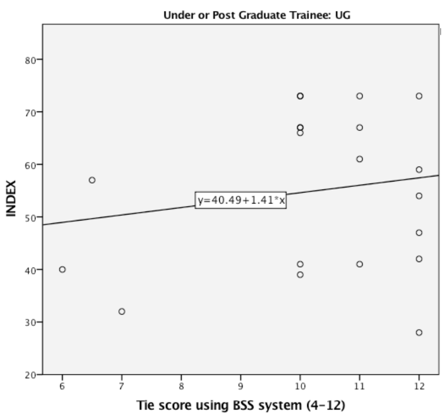


Figure 56: OSATS & Index

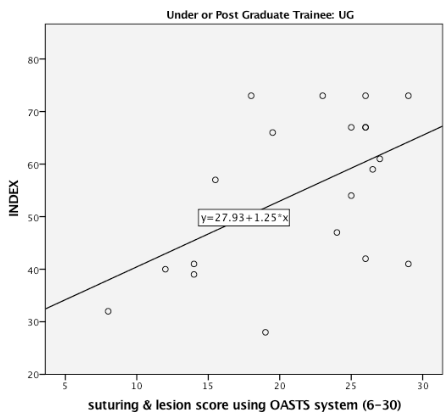


Table 29: Spearman Correlation BSS & Flying Aptitude Test Index Scores

		Hand tie Correct	Hand tie tension	Instrument tie Correct	Instrument tie tension	Suturing tissue handling	Suturing Time & Motion	Suturing Instrument handling	Lesion excision Tissue handling	Lesion excision Time & Motion	Lesion Instrument handling	Tie score using BSS system (4-12)	suturing & lesion score using OASTS system (6-30)	Total BSS score (10-42)
Psychomotor	Correlation Coefficient	-.093	-.042	.136	.565**	.128	.238	.126	.309	.458*	.451*	.046	.306	.229
	Sig. (2-tailed)	.695	.862	.568	.009	.590	.313	.596	.185	.042	.046	.847	.189	.331
	N	20	20	20	20	20	20	20	20	20	20	20	20	20
Verbal Reasoning	Correlation Coefficient	-.402	-.348	-.049	.382	.066	.044	.140	.314	.177	.261	-.283	.161	.051
	Sig. (2-tailed)	.079	.132	.836	.096	.783	.853	.557	.178	.456	.266	.227	.498	.830
	N	20	20	20	20	20	20	20	20	20	20	20	20	20
Spatial reasoning	Correlation Coefficient	.108	.087	.274	.433	.541*	.453*	.326	.398	.630**	.631**	.126	.540*	.470*
	Sig. (2-tailed)	.651	.717	.242	.056	.014	.045	.161	.082	.003	.003	.595	.014	.036
	N	20	20	20	20	20	20	20	20	20	20	20	20	20
Attentional capability	Correlation Coefficient	-.123	-.097	.037	.482*	-.042	.038	.096	.411	.199	.217	-.124	.229	.125
	Sig. (2-tailed)	.607	.685	.876	.031	.860	.874	.686	.072	.400	.358	.603	.332	.598
	N	20	20	20	20	20	20	20	20	20	20	20	20	20
Short Term Memory	Correlation Coefficient	.122	.182	-.135	.525*	.228	.467*	.328	.344	.407	.463*	.252	.382	.360
	Sig. (2-tailed)	.607	.442	.570	.017	.333	.038	.159	.137	.075	.040	.285	.096	.119
	N	20	20	20	20	20	20	20	20	20	20	20	20	20
Index Aircrew Selection Test	Correlation Coefficient	-.069	-.015	.134	.651**	.292	.381	.291	.479*	.548*	.594**	.056	.464*	.376
	Sig. (2-tailed)	.772	.951	.572	.002	.211	.097	.213	.033	.012	.006	.814	.040	.103
	N	20	20	20	20	20	20	20	20	20	20	20	20	20

Table 30: Female Flying Aptitude Test Index Scores & BSS Spearman Correlation

		Hand tie Correct	Hand tie tension	Instrument tie Correct	Instrument tie tension	Suturing tissue handling	Suturing Time & Motion	Suturing Instrument handling	Lesion excision Tissue handling	Lesion excision Time & Motion	Lesion Instrument handling	Tie score using BSS system (4-12)	suturing & lesion score using OASTS system (6-30)	Total BSS score (10-42)
Psychomotor	Correlation Coefficient	-.227	-.444	-.260	.346	.092	.274	.000	.448	.633	.633	-.490	.373	.204
	Sig. (2-tailed)	.589	.271	.534	.402	.828	.512	1.000	.266	.092	.092	.217	.363	.628
	N	8	8	8	8	8	8	8	8	8	8	8	8	8
Verbal Reasoning	Correlation Coefficient	-.861**	-.857**	-.128	-.048	-.448	-.113	-.153	.137	-.028	-.028	-.805*	-.098	-.134
	Sig. (2-tailed)	.006	.007	.762	.911	.266	.790	.718	.747	.948	.948	.016	.817	.751
	N	8	8	8	8	8	8	8	8	8	8	8	8	8
Spatial reasoning	Correlation Coefficient	-.185	-.440	.392	.024	.528	.186	-.156	.080	.679	.679	-.423	.206	.099
	Sig. (2-tailed)	.660	.276	.337	.955	.179	.659	.712	.852	.064	.064	.296	.624	.815
	N	8	8	8	8	8	8	8	8	8	8	8	8	8
Attentional capability	Correlation Coefficient	-.310	-.294	-.405	.527	-.361	.133	.107	.596	.293	.293	-.364	.239	.116
	Sig. (2-tailed)	.455	.480	.319	.180	.379	.754	.800	.119	.482	.482	.375	.568	.785
	N	8	8	8	8	8	8	8	8	8	8	8	8	8
Short Term Memory	Correlation Coefficient	-.459	-.350	-.510	.300	-.383	.013	.000	.607	.400	.400	-.581	.189	.079
	Sig. (2-tailed)	.252	.396	.197	.471	.350	.977	1.000	.110	.326	.326	.131	.654	.853
	N	8	8	8	8	8	8	8	8	8	8	8	8	8
Index Aircrew Selection Test	Correlation Coefficient	-.463	-.617	.000	.408	-.007	.205	.000	.514	.576	.576	-.603	.273	.133
	Sig. (2-tailed)	.248	.103	1.000	.316	.987	.626	1.000	.193	.135	.135	.114	.513	.754
	N	8	8	8	8	8	8	8	8	8	8	8	8	8

Table 31: Male Flying Aptitude Test Index Scores & BSS Spearman Correlation

		Hand tie Correct	Hand tie tension	Instrument tie Correct	Instrument tie tension	Suturing tissue handling	Suturing Time & Motion	Suturing Instrument handling	Lesion excision Tissue handling	Lesion excision Time & Motion	Lesion Instrument handling	Tie score using BSS system (4-12)	suturing & lesion score using OASTS system (6-30)	Total BSS score (10-42)
Psychomotor	Correlation Coefficient	.010	.307	.356	.782**	.587*	.419	.461	.474	.696*	.665*	.378	.616*	.570
	Sig. (2-tailed)	.975	.332	.257	.003	.045	.176	.132	.120	.012	.018	.226	.033	.053
	N	12	12	12	12	12	12	12	12	12	12	12	12	12
Verbal Reasoning	Correlation Coefficient	-.144	-.008	.046	.589*	.229	.105	.233	.360	.248	.355	.012	.235	.176
	Sig. (2-tailed)	.655	.980	.888	.044	.473	.745	.467	.250	.438	.257	.971	.463	.584
	N	12	12	12	12	12	12	12	12	12	12	12	12	12
Spatial reasoning	Correlation Coefficient	.303	.379	.313	.681*	.498	.535	.480	.485	.537	.556	.437	.605*	.576*
	Sig. (2-tailed)	.339	.224	.322	.015	.099	.073	.114	.110	.072	.060	.156	.037	.050
	N	12	12	12	12	12	12	12	12	12	12	12	12	12
Attentional capability	Correlation Coefficient	-.040	.035	.354	.467	.172	.065	.178	.424	.177	.258	.017	.288	.218
	Sig. (2-tailed)	.902	.915	.259	.126	.592	.840	.580	.170	.582	.418	.959	.364	.495
	N	12	12	12	12	12	12	12	12	12	12	12	12	12
Short Term Memory	Correlation Coefficient	.509	.614*	.354	.648*	.519	.726**	.484	.158	.487	.504	.750**	.476	.579*
	Sig. (2-tailed)	.091	.034	.259	.023	.084	.008	.111	.623	.108	.095	.005	.118	.049
	N	12	12	12	12	12	12	12	12	12	12	12	12	12
Index Aircrew Selection Test	Correlation Coefficient	.198	.392	.308	.825**	.586*	.539	.514	.499	.591*	.663*	.450	.629*	.601*
	Sig. (2-tailed)	.536	.207	.330	.001	.045	.071	.088	.098	.043	.019	.142	.028	.039
	N	12	12	12	12	12	12	12	12	12	12	12	12	12

Even though females had higher mean scores for all the ten BSS Tests (Table 28) it had not shown statistical significance. However, when running the Spearman correlation between the BSS tests and the Flying Aptitude Test domains, the results for female participants did not show any statistical significance (Table 30) whilst the male (Table 31) participants had a positive correlation between the Flying Aptitude Test Index Scores and OSATS scores (r=0.629; p=0.028) and with the total BSS scores (r=0.601; p=0.039). The scatter plot shows a gradient difference between genders (Figure 57, Figure 58).

Figure 57: Gender difference BSS & Flying Aptitude Test Index scores scatter plot

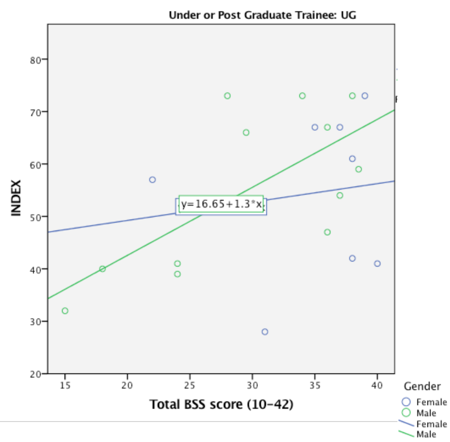
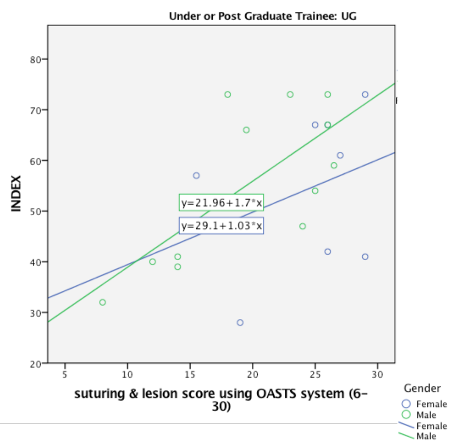


Figure 58: Gender difference OSATS & Flying Aptitude Test Index scores



No statistical significance was found in the results between the genders when the two correlation coefficients in the two independent samples were analysed using the Fisher transformation test (Table 32).

Table 32: Fisher Transformation of Correlation Coefficients between genders (F.A.T & BSS)

Gender	Correlation Coefficient	Fisher r to z transformation	Significance	
♀ n = 78	-0.288	z = -0.85	0.1977	1 tailed
♂ n = 75	-0.155		0.3953	2 tailed

4.5 Computer Games & Flying Aptitude Tests Results

Computer game use results were available on 168 participants. 136 (81%) of those who responded (n=168) replied yes to playing computer games (55 females and 81 males Table 33). 95.3% of male responders played computer games and less than 10% described themselves as beginners when 71% of female responders described themselves as beginners (out of the 66% who responded yes to playing computer games) (Table 34). The difference between the genders in the use of computer games (female 66.3% vs. 96.3% $p<0.001$) and their self-reported ability was statistically significant ($p<0.011$).

Table 33: Playing Computer Games

	All Participants				Female				Male			
	ALL		UG		ALL		UG		ALL		UG	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Yes	136	80.95	120	80	55	66.27	51	64.56	81	95.29	69	97.18
No	32	19.05	30	20	28	33.73	28	35.44	4	4.71	2	2.82
Total	168	100	150	100	83	100	79	100	85	100	71	100

Table 34: Computer Games Playing Levels

	Total		Female		Male	
	Freq.	%	Freq.	%	Freq.	%
Beginner	47	34.6	39	70.9	8	9.9
Average	54	39.7	14	25.5	40	49.4
Intermediate	26	19.1	2	3.6	24	29.6
Expert	9	6.6	0	0	9	11.1
Total	136	100	55	100	81	100
Missing	107		61		46	
Total	243		116		127	

Data on Computer Games playing who had undertaken the Flying Aptitude Test was available in 162 participants of whom 133 (82%) answered yes to computer games of whom 117 were undergraduates. A third of undergraduate female participants did not play computer games compared to 98.5% of the male participants who did play computer games (Table 35).

Table 35: Computer game frequency

	Computer Games		
FLYING APTITUDE TEST	yes	no	total
ALL	133 (82.1%)	29 (17.9%)	162
UG	117 (80.7%)	28 (19.3%)	145
UG Male	67 (98.5%)	1(1.5%)	68
UG Female	50 (64.9%)	27 (35.0%)	77

Table 36: Undergraduate Comparison of Flying Aptitude Test Domains Means (Independent t-test) between Computer Game players and non-players

	Computer Games	N	Mean	SD	Std. Error Mean	Sig. (2-tailed)
Psychomotor	Yes	117	4.17	1.802	0.167	
	No	28	2.54	1.527	0.289	<0.001
Verbal Reasoning	Yes	117	5.88	1.848	0.171	
	No	28	5.71	1.922	0.363	0672
Spatial reasoning	Yes	117	4.81	1.786	0.165	
	No	28	5	1.743	0.329	0.616
Attentional capability	Yes	117	4.27	2.144	0.198	
	No	28	3.29	2.386	0.451	0.034
Short Term Memory	Yes	117	4.76	1.959	0.181	
	No	28	4.57	2.008	0.379	0.648
INDEX Flying Aptitude Test	Yes	117	52.38	13.463	1.245	
	No	28	43.79	13.178	2.49	0.003

The Flying Aptitude Test Index Scores mean difference between computer game players and non-players was 52.38 vs 43.79 p=0.003 (Figure 59). Out of the five Flying Aptitude Test domains the Psychomotor domain (Figure 60) was the aptitude with the largest difference in Mean values between participants who played computer games and those who didn't (Yes=4.17 vs. No=2.54 p<0001) (Table 36).

Figure 59: Flying Aptitude Test Index score difference between computer gamers and non-gamers

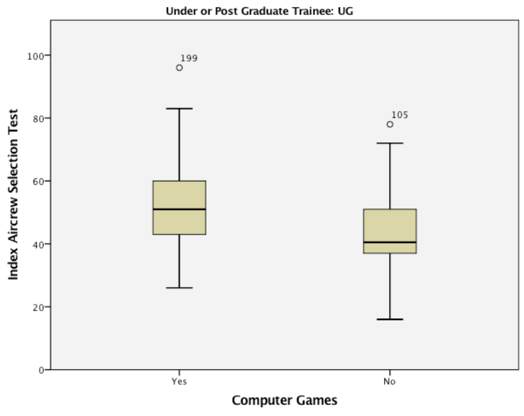
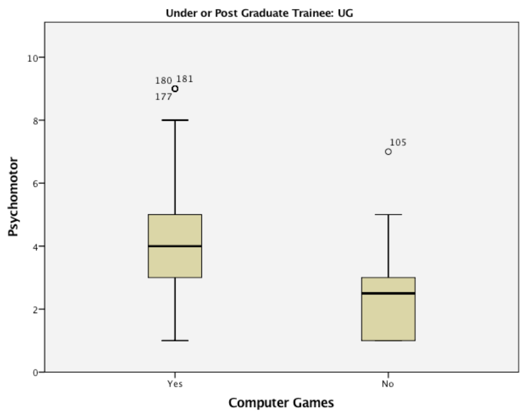


Figure 60: Psychomotor aptitude differences between computer gamers and non-gamers



There were only 4 out of 85 male participants who had participated in the Flying Aptitude Test who did not play computer games, compared to a third (27 out of 81) of female participants who admitted to not playing any computer games. For the female group the psychomotor aptitude also showed a difference in the mean values which was statistically significant (Yes 3.20 vs. No 2.44; p=0.024) (Table 37).

Table 37: Independent T-Test computer games in female participants and Flying Aptitude Test Scores.

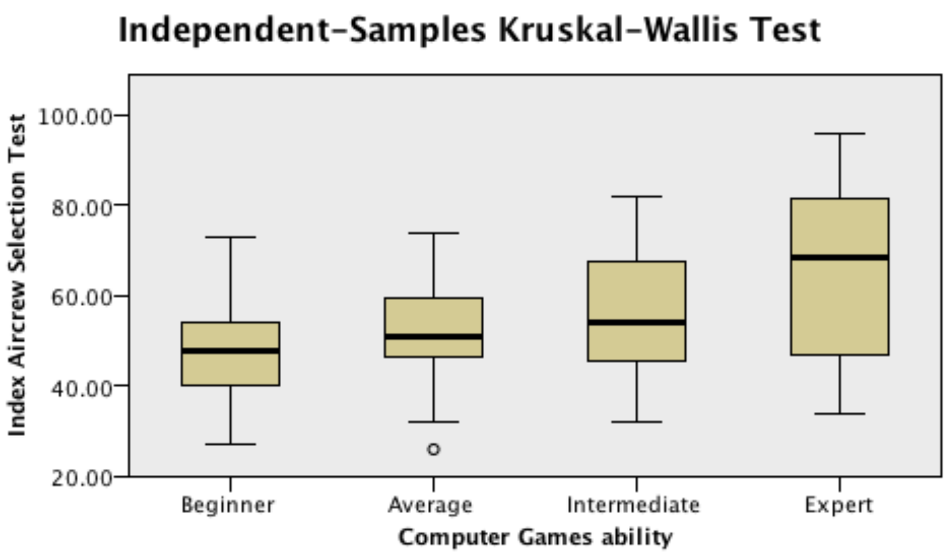
FEMALE	Computer Games	N	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Psychomotor	Yes	54	3.2	1.155	0.157	
	No	27	2.44	1.476	0.284	0.024
Verbal Reasoning	Yes	54	6.13	2.065	0.281	
	No	27	5.7	1.958	0.377	0.369
Spatial reasoning	Yes	54	5	1.59	0.216	
	No	27	4.96	1.765	0.34	0.927
Attentional capability	Yes	54	3.57	2.006	0.273	
	No	27	3.22	2.407	0.463	0.516
Short Term Memory	Yes	54	5.04	1.832	0.249	
	No	27	4.56	2.044	0.393	0.307
INDEX Flying Aptitude Test	Yes	54	48.61	11.095	1.51	
	No	27	43.19	13.033	2.508	0.07

The difference in the aptitudes was even more noticeable when comparing the different levels of computer game ability. The number of females reporting as beginner was 70.4% compared to only 10.1% of males (Table 35).

Table 38: Computer Game Playing Level in Flying Aptitude Test participants divided by gender.

	Total		Female		Male	
	Freq.	%	Freq.	%	Freq.	%
Beginner	46	34.6	38	70.4	8	10.1
Average	53	39.8	14	25.9	39	49.4
Intermediate	25	18.8	2	3.7	23	29.1
Expert	9	6.8	0	0	9	11.4
Total	133	100	54	100	79	100
Missing	97		58		39	
Total	230		112		118	

Figure 61: Flying Aptitude Test Index score differences between Computer Games Ability in UG



Total N	117
Test Statistic	9.680
Degrees of Freedom	3
Asymptotic Sig. (2-sided test)	.021

The Kruskal-Wallis Test boxplot (Figure 61) shows the difference in the undergraduate's Flying Aptitude Test Index scores between the four computer games ability groups. This was statistically significant at p=0.021.

In all five aptitudes and the total Index scores the Mean values for the expert group was higher than the beginner group (Table 36).

The Psychomotor Mean value for the Expert gamers was almost twice that of the Beginners (Expert 6.44 vs. 3.37 Beginner; p=<0.014). Attentional capability also showed a statistically significant difference (Expert 5.78 vs. 3.39 Beginner; p=<0.017). The overall Index score difference was greater than 20% (Expert 67.89 vs. 47.61 Beginner; p=0.021) (Figure 61).

Table 39: Flying Aptitude Test Means Comparison (Independent t-test) between beginner and expert gamers

	Computer Games ability	N	Mean	SD	Std. Error Mean	Sig. (2-tailed)
Psychomotor	Beginner	46	3.37	1.306	0.193	
	Expert	9	6.44	2.351	0.784	<0.014
Verbal Reasoning	Beginner	46	5.91	1.998	0.295	
	Expert	9	6.44	2.297	0.766	0.531
Spatial reasoning	Beginner	46	4.78	1.504	0.222	
	Expert	9	5.33	1.803	0.601	0.409
Attentional capability	Beginner	46	3.39	1.96	0.289	
	Expert	9	5.78	1.986	0.662	<0.017
Short Term Memory	Beginner	46	4.52	1.761	0.26	
	Expert	9	5.56	2.128	0.709	0.2
INDEX Flying Aptitude Test	Beginner	46	47.61	11.125	1.64	
	Expert	9	67.89	21.233	7.078	0.021

The four different abilities were grouped into two: A: Beginner/Average (n=99) and B: Intermediate/Expert (sum n=34). Independent t-test comparison of Means showed that in all Aptitudes the Intermediate/Expert group had a higher Mean score than the Beginner/Average group with Psychomotor (A=3.99 vs. B=5.38; p=<0.012), Attentional Capability (A=3.96 vs. B-5.12; p=<0.019) and Index scores (A=50.96 vs. B=60.5; p=<0.016) having statistical significance.

There was no statistical significant difference between female and male's Flying Aptitude Test Mean scores in the computer games beginner group. When comparing the Average level group there was a statistical difference between the genders in the Psychomotor (female 3.21 vs. male 4.13 p<0.001) and Short Term Memory (female 4.61 vs male 4.13 p=<0.016) aptitudes (Table 37). These differences between genders were smaller than when compared regardless of computer game playing levels.

Table 40: Average level of computer gamers' Flying Aptitude Tests Mean Comparison (Independent t-test) in the between the genders.

Gaming Level AVERAGE	Gender	N	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Psychomotor	Female	38	3.21	1.212	0.197	
	Male	8	4.13	1.553	0.549	<0.001
Verbal Reasoning	Female	38	6	2.105	0.342	
	Male	8	5.5	1.414	0.5	0.088
Spatial reasoning	Female	38	4.89	1.485	0.241	
	Male	8	4.25	1.581	0.559	0.455
Attentional capability	Female	38	3.37	2.123	0.344	
	Male	8	3.5	0.926	0.327	0.682
Short Term Memory	Female	38	4.61	1.779	0.289	
	Male	8	4.13	1.727	0.611	<0.016
Index Flying Aptitude Test	Female	38	47.42	11.479	1.862	
	Male	8	48.5	9.899	3.5	0.387

4.6 Interest in a Career in Surgery or Gastroenterology

Data was available in 67.9% (n=165) of the participants and those interested in a career in surgery accounted for 58.8% (n=97); those not interested 38.2% (n=63); 1.2% (n=2) maybe and 1.8% (n=3) unknown. Interest in surgery had less than ten per cent difference between the genders (female 54.2% vs. 63.4% male) and this was not statistically significant (Independent sample t-test p=0.904).

Interest in a career in Gastroenterology was expressed by 55 participants (33.3%); not interested by n=64 (38.8%); n=4 (2.4%) answered maybe and n=42 (25.5%) did not know. There was a 16.2% difference between the genders but this again was not statistically significant.

The level of interest in either Surgery or Gastroenterology for the 19 participants whose English was not their first language did not differ to the rest of the 146 participants who were native English speakers (Independent sample t-test p=0.624 & p=0.51) (Table 41).

Table 41: Interest in a career in Surgery or Gastroenterology by English being the first or second language

N=19	English	N	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Surgery	1st	146	1.45	0.622	0.052	0.624
	2nd	19	1.53	0.612	0.14	
Gastroenterology	1st	146	2.22	1.177	0.097	0.519
	2nd	19	2.05	1.026	0.235	

A total of 159 participants recorded on their preference for a career in surgery and a career in gastroenterology.

Table 42: Interest in a Career in Surgery or Gastroenterology in the Flying Aptitude Test participants

Interest in a Career in	Surgery		Gastroenterology	
	Freq.	%	Freq.	%
Yes	93	58.5	53	33.3
No	61	38.4	61	38.4
Maybe	2	1.3	4	2.5
Unknown	3	1.9	41	25.8
Total	159	100	159	100
Missing	71		71	
Total	230		230	

93 out of the 159 (58.5%) of the participants expressed an interest in a career in surgery (Table 42). For the five Flying Aptitude Tests and the final Index scores, Independent t-test comparison of Means showed no difference between the participants who expressed an interest in a career in Surgery to those who did not (Table 39). Of the 159 participants whose interest in gastroenterology was recorded, 38% were not interested, 33% said yes, 26% did not know and 3% chose maybe (Table 42).

There was no statistical difference between the means of the five Flying Aptitude Test as well as the final Index score for those participants who showed an interest in a career in gastroenterology and those who didn't.

There were no statistical differences in any of the Lap Sim Results between the participants who showed an interest in surgery (Yes n=96; No n=61) or gastroenterology (Yes n=55; Non=64) using the Independent t-test for Means.

Out of the 26 who participated in the BSS, 17 participants responded if they had an interest in a career in surgery; of which eleven (64.7%) had shown an interest, five responded no and one maybe (Table 43)

Table 43: Frequency tables for BSS participants who have an interest in a career in Surgery or Gastroenterology

Interest in a career in	Surgery		Gastroenterology	
	Freq.	%	Freq.	%
Yes	11	64.7	5	29.4
No	5	29.4	7	41.2
Maybe	1	5.9	5	29.4
Total	17	100	17	100
Missing	9		9	
Total	26		26	

Only five participants showed an interest in a career in gastroenterology, all of whom were medical students (29.4%), seven participants (41.2%) answered ‘no’ and five ‘maybe’ (Table 43). Five medical students, one FY1 and one CT1 trainee were in the ‘no’ group, the CT1 trainee being the only one in a surgical training path.

4.7 Self Rating of prior to the Flying Aptitude Test

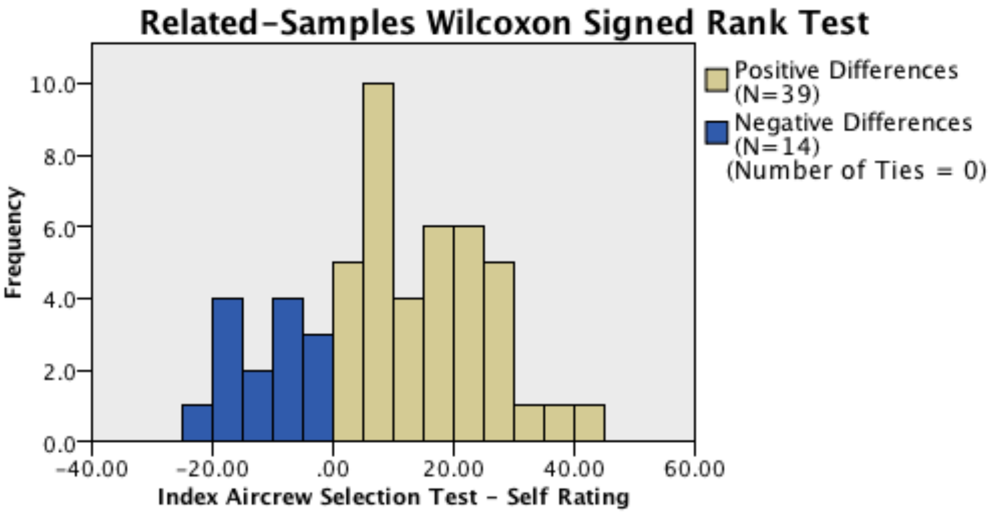
Self-Rating Questionnaire data from 111 undergraduate participants was collected prior to undertaking the Flying Aptitude Test at the OASC, RAF Cranwell. The mean percentage was 47.90% (min=1.71% to max 100%) (Table 44).

Table 44: Self Rating frequency table

Self-Rating	All Participants	Female	Male
N	111	53	58
Mean	47.90	37.23	57.65
Std. Error of Mean	1.71	2.19	1.84
95 % CI for Mean Lower	44.50	32.85	53.96
95 % CI for Mean Upper	5<0.01	41.61	61.33
Median	5<0.01	34.70	57.00
Mode	48.00	26.00	61.30
Std. Deviation	18.06	15.91	14.00
Variance	326.17	253.11	196.07
IQR	27.40	23.35	15.47
Skewness	-0.15	-0.01	0.07
Kurtosis	-0.05	-0.31	0.92
Range	96.70	74.00	75.30
Minimum	3.30	3.30	24.70
Maximum	10<0.01	77.30	10<0.01
Percentiles 25	33.30	25.65	5<0.01
50	5<0.01	34.70	57.00
75	60.70	49.00	65.48

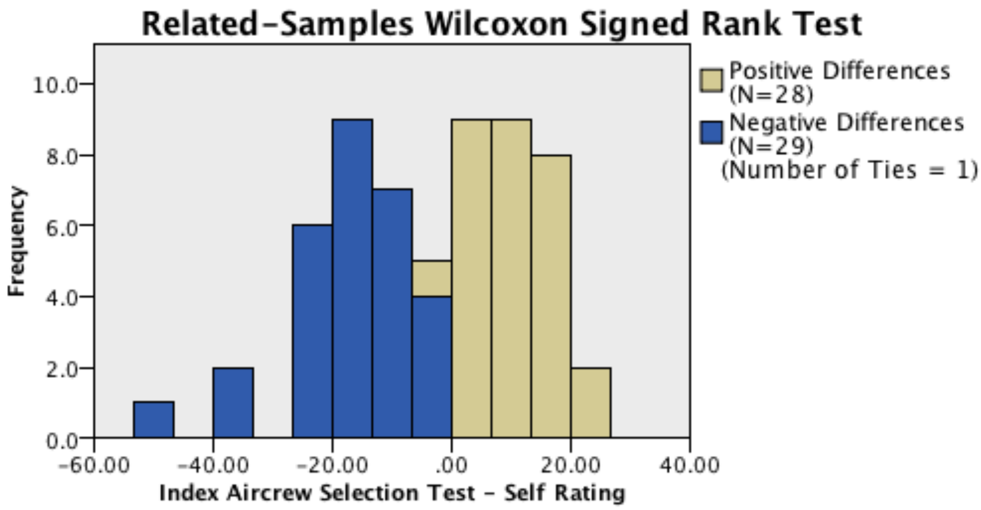
Wilcoxon Signed Rank Test was used to on the Self-Rating data for both female and male undergraduates. 39 out of the 53 female undergraduates scored higher ($p<0.011$) in their Flying Aptitude Test than their Self-Rating scores. Male participants did not show a statistically significant difference between their Self-Rating score and the Flying Aptitude Test results.

Figure 62: Self-Rating vs Flying Aptitude Test Index Scores: Wilcoxon Signed Rank Test for Female Undergraduates



Total N	53
Test Statistic	1,113.500
Standard Error	112.953
Standardized Test Statistic	3.524
Asymptotic Sig. (2-sided test)	.000

Figure 63: Self-Rating vs Flying Aptitude Test Index Scores: Wilcoxon Signed Rank Test for Male Undergraduates



Total N	58
Test Statistic	669.000
Standard Error	125.849
Standardized Test Statistic	-1.251
Asymptotic Sig. (2-sided test)	.211

Table 45: Self Rating Correlation with Flying Aptitude Tests

SELF RATING n=124	Pearson Correlation	Sig. (2-tailed)
Psychomotor	0.511**	<0.001
Perceptual	0.09	0.319
Spatial reasoning	0.02	0.829
Attentional capability	0.253**	<0.015
Short Term Memory	0.079	0.382
Index Flying Aptitude Test	0.390**	<0.001
** Correlation is significant at the 0.01 level (2-tailed).		

The Self Rating percentage scores correlated positively with the Psychomotor (r=0.511, p<0.001) and Attentional Capability Aptitudes (r=0.253, p=<0.015) which resulted in a positive correlation with the Index score (r=0.390, p<0.001) (Table 45).

The only variable that showed a statistical difference was between participants who played computer games and those who didn't. Gamers had a mean score of 51.23% compared to non-gamers 36.37%, p=<0.012 (Table 46). None of the other variables (training grade, English as a first language, Chopstick use, age, interest in a career in surgery or gastroenterology) had shown any statistical difference .

Table 46: Self Rating Mean Difference between computer gamers and non-gamers

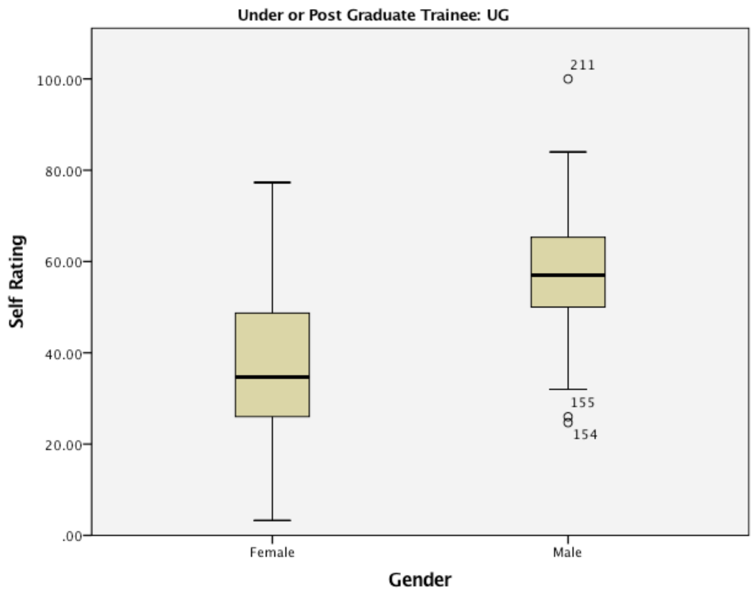
Computer Games	N	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Yes	82	51.2268	17.40382	1.92193	
No	22	36.2682	18.94899	4.03994	<0.012

There was a statistically significant difference between what female and male participants rated on how good their Flying Aptitude Test result would yield. The difference was greater than twenty percentage points (female 36.4% vs males 57.3%, p<0.001) (Table 47). All of the participants in the top 10% self-rating scores were males and from the lowest 10% all were female but one (Figure 45).

Table 47: Self Rating Mean Difference between Genders

Gender	N	Mean	Std. Deviation	Std. Error Mean	Sig. (2-tailed)
Female	56	36.3625	16.72244	2.23463	
Male	68	57.325	15.19756	1.84297	<0.001

Figure 64: Self Rating Gender Differences



Both female and male participant's Self Rating scores correlated significantly with the Psychomotor Aptitude. Female participants not only had a greater correlation value on the Psychomotor Aptitude (r=0.378, p=<0.014 vs r=0.257, p=0.034) but also had a statistically significant correlation with the Verbal Reasoning (r=0.316, p=0.018) (Table 48).

Table 48: Gender differences in Self Rating Correlation with Flying Aptitude Tests

	FEMALE n=56		MALE n=68	
	Pearson Correlation	Sig. (2-tailed)	Pearson Correlation	Sig. (2-tailed)
Psychomotor	0.378**	<0.014	0.257*	0.034
Verbal Reasoning	0.316*	0.018	0.029	0.812
Spatial reasoning	0.043	0.755	0.135	0.272
Attentional capability	0.152	0.263	0.186	0.128
Short Term Memory	0.158	0.245	0.121	0.324
Index Flying Aptitude Test	0.378**	<0.014	0.238	0.051

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

4.8 Feedback Questionnaire on the Aptitude Test

From the Participant's Feedback form (Appendix 10), the result to the question on whether a Computerised Aptitude Test such as the Flying Aptitude Test should be introduced into UK surgical selection was analysed (Table 49).

More than two thirds of participants (64.5%) responded positively to the introduction of such Aptitude Test into Surgical Selection in the UK (Strongly Agree 32.9%, n=25 and Agree 31.6%, n=24). 27.6%, n=21 were neutral to the idea and only six participants (7.9%) disagreed. There were no participants who Strongly Disagreed.

Table 49: Do you agree with an Aptitude Test for the Selection of Surgical Trainees?

N=76	Frequency	Percent	Cumulative Percent
Strongly Disagree	0	0	0
Disagree	6	7.9	7.9
Neutral	21	27.6	35.5
Agree	24	31.6	67.1
Strongly Agree	25	32.9	100
Total	76	100	

Chapter 5: Discussion

"Achievement is talent plus preparation"

Malcolm Gladwell, Outliers: The Story of Success[119].

5.1 Flying Aptitude Test & Laparoscopic Simulator Tests

The participants in our study scored comparatively to the Royal Air Force (RAF) candidates on all of the five Flying Aptitude Tests (Psychomotor, Perceptual, Spatial, Attentional & Short Term Memory), the highest domain being Verbal Reasoning and a mean total Index score of 51.78%. The medical participants in our study compared favourably to the RAF candidates who have undergone the RAF Selection Test at the Officers and Aircrew Selection Centre within the same year. Every single of the five Flying Aptitude Test domains correlated with each other significantly.

The Flying Aptitude Test correlated with the Total Laparoscopic Simulation (Lap Sim) Time at r=-0.275; p=0.001. The higher the Flying Aptitude Test Index, the faster the completion of the four laparoscopic simulated tasks. The total Lap Sim Time was the sum of the Bean Drop, Block Move, Appendicectomy and Common Bile Duct Cannulation Tasks.

The highest correlation value was between the Psychomotor Aptitude and the Lap Sim Tests times at r=0.300; p<0.001. The higher the Psychomotor

Aptitude Test score the faster the time to complete the Lap Sim tasks. The correlation for the undergraduate participants was r = -0.275, p<0.001. Correlation with the individual Lap Sim tests was also positive for the Bean Drop (r=-0.270, p<0.001), Block Move (r=-0.186, p=<0.021), and Appendicectomy (r=-0.187, p=<0.021) Tests.

Attentional Capability aptitude was found to have a statistically significant correlation with the three of the four Lap Sim tests. The two basic tests of Bean Drop (r=-0.239, p=<0.003) and Block move (r=-0.221, p=<0.006) and the clinically orientated Bile Duct Cannulation Test (r=-0.188, p=<0.02). Even though the Appendicectomy task did not show significant correlation the total Lap Sim Time still showed a statistically significant negative correlation of r=-235, p=-.003.

The first hypothesis can be accepted as the Flying Aptitude Test Index score does correlate with the Total Lap Sim Time. The Psychomotor and the Attentional Capability domains correlate significantly as well. The caveat explained later is that this hypothesis only applies to female participants as for the male participants none of the four Lap Sim tasks correlated with any of the Flying Aptitude Test domains.

As the RAF at the OASC believe, previous flying experience will not affect the performance at the Flying Aptitude Test as it is innate skills that the tests are measuring. The third hypothesis was accepted that there was no difference between undergraduates and postgraduates with varying levels of previous surgical experience in their Flying Aptitude Test results. 199 undergraduates'

Flying Aptitude Test was correlated with 31 postgraduate participants' results and there was no statistically significant difference in any of the aptitudes domain. The mean Flying Aptitude Test Index score was 51.64 in undergraduates and 52.71 with p=0.741.

5.2 Flying Aptitude Test & Basic Surgical Skills Tests

The BSS Test total score is the sum of ten tests with two scoring systems. The 3 point scale BSS scoring Suture tying tests (hand and instrument) and the 5 point scale OSATS scoring 'Suturing' and Excision tests.

The Flying Aptitude Test Index and the total BSS scores showed a positive correlation of r = 0.376 (Figure 54), but the p value was outside of significance p=0.103. The reason may be due to the non-existent correlation between the 3 point scale Tying Tests ('Hand Tie' and 'Instrument Tie') and the Flying Aptitude Test INDEX result (r=0.056; p=0.814) (Figure). The correlation between the total scores of both of the five-point OSATS scale tests (the 'Suturing' and 'Lesion Excision' Tests) showed a positive correlation with the Flying Aptitude Test Index score at r=0.464; p=0.04.

Spatial Reasoning did not show a significant correlation with the Lap Sim Test results but had the highest correlation with the BSS tests. Spatial Reasoning had the highest correlation with the BSS total score (r=0.470, p<0.036) and the highest OSATS score correlation of r=0.540; p=0.14. Spatial Reasoning provided the highest correlation between all ten BSS scores and the five Flying Aptitude Test domains (r=0.631; p=003).

The second hypothesis where the Flying Aptitude Test results will correlate positively with the Basic Surgical Skills Tests results can be accepted for the more advanced surgical skills ('Suturing' and 'Lesion Excision') and the 5-point OSATS system used for scoring rather than the cruder 3 point Basic Surgical Skills official course measurements. Inversely to the Lap Sim Tests, the correlation between the open surgical BSS tests was only statistically significant in the male participant group.

5.3 Gender Differences

The Psychomotor Domain showed the greatest statistical difference in the distribution of scores between the genders (Female 2.93 vs. 5.07 Stanine Scale 1-9) and as it was the heaviest weighed aptitude at 40%, it resulted in the Flying Aptitude Test Index score to have a 10.6% difference (female 46.51 vs. male 5.14; $p<0.001$). Male participants also had a higher mean on the Attentional Capability. Female participants scored higher in the Verbal Reasoning but this was not statistically significant.

Male participants completed the Lap Sim Tests in a shorter period of time compared to the female participants (Total Lap Sim Time female 800 seconds vs 674 seconds male $p=0.01$), but correlation analysis between the Flying Aptitude Test domains & Lap Sim Tests only showed statistically significant correlations with the Psychomotor, Attentional Capability and Total Lap Sim ($r=-0.246$, $p=0.027$) in female participants. Analyses in male participants did not show any correlation of the data. The first hypothesis of this study can only relate to the female participants.

The number of participants in the Basic Surgical Skills tests were small but in every single of the ten sets of data the mean scores for the female participants (n=9) was higher than the males (n=13). The results were not statistically significant but the box plot and SD data showed the male participants having a much wider distribution compared to the female participants.

It may potentially be due to the small number of participants but correlation of the BSS scores with the Flying Aptitude Test Domains showed that only in male participants there was any significance to be found. The male participants had a positive correlation was found between the Flying Aptitude Test Index and both the OSATS scores (r=0.629; p=0.028) and with the total BSS scores (r=0.601; p=0.039).

The latest GMC report on 'The state of medical education and practice in the UK'[120] states that in 2013 female doctors made up 44% of licensed doctors and 54% of medical students were female but only 10% of consultant surgeons were female.

Our study had similar number of female (47.7%) and male (52.3%) participants and even thought not statistically significant, there was a difference between the genders in their interest in surgery; 54.2% of females and 63.4% male participants showed an interest in a career in surgery. The difference between the genders was similar to a study by Fitzgerald[121] where newly qualified graduates from the University of Nottingham Medical School (the similar cohort as our study) found that out of two hundred and eight questionnaires the male respondents were significantly more likely to rate surgery as an attractive or very attractive career. Only 23% of female junior doctors expressed an interest in a career in surgery as opposed to 42% of male junior doctors. The difference was nearly twice as much than in our study which can be attributed to most of our cohort being studied during their surgical attachment as an undergraduate. In Fitzgerald's study, irrespective of career interests, 59% of male and 68% of female respondents believed surgery was not a career welcoming women, stating reasons such as difficulty maintaining family life, limited flexible training, and lack of role models.

This gender disparity is not only a problem of the UK but found to be true in many other countries. This ratio has also been described in the US where women represent 15% of practicing general surgeons and a study by Bruce[122] from Washington DC concludes that the majority of the 334 responses which included medical students, residents and practicing physicians indicated perceived gender-based discrimination during training and practice. This gender-based discrimination was perceived to come from both sexes and had a significant impact on women surgeons. A study of 605 medical students from five medical universities in Poland also showed a difference between the genders, where 15% of women and 30% of men declaring to pursue a surgical specialisation[123]. A cross sectional study of 271 interns at the Ibadan

University College Hospital, Nigeria found that males were more likely to choose surgery than females (52.1% vs. 13.0%, p < <0.011)[124].

UK Surgical training takes a minimum of 10 years after graduation of medical school, and the number of female consultant surgeons will eventually improve as the number of female surgical trainees has risen to 28% in 2013[125]. A presentation at the American College of Surgeons 95th Annual Clinical Congress in San Francisco by Elisabeth Davis[126] described how the gender gap among US medical graduates entering general surgical training appeared to be closing. Between the academic years of 1999-2000 and 2004-2005 the percentage of US medical graduates applying to General Surgical programs increased from 27% to 33% and those entering General surgical residencies from 32% to 40%.

M. Baily, chief Psychologist from the Royal Air Force Selection Centre, RAF Cranwell[127] explained how this 10% difference between the genders found in this study is present also between the female and male candidates who wish to join the RAF and undergo the Flying Aptitude Test. Bayley explained that this gap between the genders has been reducing over the last half century since the aptitude test has been adopted into selection. When the aptitude test was first introduced during WWII the way that boys and girls played in the UK differed significantly to Millennials. Boys would spend a lot more time outdoors climbing trees and girls would be encouraged to help with household chores. Shortening of the gap between the genders supports the theory that the psychomotor ability is due to early developmental exposure rather than genetic predisposition as most modern girls will have the similar opportunities in play as boys.

The Self Rating questionnaire was conducted to evaluate insight on the participants' own level of skill. The perceived difference between the genders was noted when the Self Rating Questionnaire prior to the Flying Aptitude Test showed a statistically significant >20% difference. Female participants were rating themselves at only 36.4% whilst male participants rated themselves at 57.3%; $p<0.001$). The correlation was significant in the Psychomotor Aptitude in both genders but Female participants had a greater correlation value as well as correlating positively in the Verbal Reasoning.

The hypothesis that female and male participants would perform similarly cannot be accepted as the findings from this study has revealed how different the results can be in every single simulated environment.

5.4 Effect of Computer Games Play

The availability on the selection of different ways to play did not seem to have resulted in the same interest between the genders. There was a clear difference in the percentage of computer gamers between the sexes and on how they rated themselves when asked about their level of proficiency. 95.3% all of the male participants played computer games compared (97.2% in the undergraduate group) to only two thirds (66.3%) of the female participants, of whom the majority described themselves as beginner players. Players who described themselves as experts were all male.

When comparing the mean values of the Flying Aptitude Test results between the computer game players and non-players, the Psychomotor aptitude showed the greatest difference (gamers 4.35 vs 2.62 non gamers, $p<0.001$) and this difference was present when comparing female gamers and non-gamers (3.2 vs.2.44, $p=0.024$).

A questionnaire specifying hours of computer game play per day or per week grouped may have provided a more objective measure and data could have been analysed as a continuous variable instead of an ordinal variable as it is in this study. But self-reporting of hours played is also known to be very subjective.

This study is unable to explain if those participants who play computer games do so because they have an innate Psychomotor and Attentional Capability abilities that gives them an advantage at playing computer games or playing computer games improves their Psychomotor and Attentional Capability Aptitudes. What we can conclude is that there is a clear correlation between them and playing computer games results in higher aptitude scores and female participants play less and at a lower level than the male participants.

A study by Beermann[128] had shown that three-dimensional (3D) visualisation improved the understanding of surgical liver anatomy compared to students in the two dimensional (2D) group, and that this improvement was statistically greater in male students. But our study did not show a statistically significant difference in Spatial Reasoning between the genders (female 4.95 vs 5.01 male; p=0.819).

Various papers have already been published the relationship between the use of computer games for improvement on laparoscopic simulated tasks and even a custom-made motion-sensitive video game is being developed that uses Wii Remote add-ons that claims to specifically be aimed at training surgical skills [129]. A review paper in the International Journal of Surgery[130] concluded that even though there is some evidence on the relationship between laparoscopic skills and the use of computer games and that it may be due to better psycho-motor skills in gamers, further research would be useful to demonstrate whether there is a direct transfer of skills from laparoscopic simulators to the operating table.

The paper on the American Journal of Surgery [131] also suggests that the use of computer games should not only be for the novice surgeons but as a warm up prior to surgery.

This study adds to the body of evidence that shows a clear advantage by the computer gamers in all of the surgical simulation tests and the big difference on computer gaming proportion and level in the female participant group may be the reason for the difference in their simulated scores.

5.5 Nature vs. Nurture

There are two schools of thought on the ability of a trainee to be able to achieve mastery. The first is that as long as each person is given sufficient time to develop, everybody is trainable. The 1931 Theory of Education in the United States by Albert Jay Nock[132] championed the egalitarian belief of universal education. Training will produce improvement in anyone, even though the amount of improvement was dependant on the individual and time taken training being a mere matter of money, equipment and specific instruction. The other school of thought believes different people possess different levels of aptitude suitable for specific professions especially the high-stakes ones such as surgery or aviation, which might not be suited for everyone.

A publication by Ericsson, Krampe and Tesch-Romer[133] whose 1993 paper coined the 10,000 hours of practice that makes an expert, suggest that practice effectively eliminates the role of pre-existing individual differences in abilities in determining task performance, whilst Ackerman[56] and Matthews[134] support the notion that general ability plays an important role in predicting individual differences in task performance even after extended task practice. Ackerman[135] in his 2000 paper proved to evaluate whether individual differences in performance after practice could be predicted from pre-existing ability measurements. His results showed that not only general ability influenced the outcome but the investigation findings included that even though the magnitude of the differences declined with practice between the participants, the performance variables only declined marginally and the difference was as well predicted after practice as it was at the beginning of practice. In the same paper Ackerman tests a touch-panel computerised installation for the measurement of psychomotor abilities prior to practice and shown a significant incremental predictive validity for the final complex skilled task (in this case the Air Traffic Control criterion task). Ackerman supports the claim made by Adams in 1957 that general intelligence, psychomotor and processing speed measurements can be used for the prediction of skill task performance when administered prior to task practice, even when practice made the overall variance between subjects decline.

The New York Times Best Seller, author, journalist and speaker Malcolm Gladwell mentions how genius is not the only or even the most important thing when determining a person's success. Throughout his book ‘Outliers’ he emphasises that the innate talent or intelligence requires the right environment and support for the individual to be able to invest in the rule of the 10,000 hours by Ericsson[60] leading to success. Even though his take home message is how society can provide the right environment and opportunities for success to occur, he does not deny the existence of innate talent or aptitude.

Surgeon and educator and founder of the Mayo Clinic, USA, Charles Mayo believed that experience does not necessarily equate to competence: he quotes that experience can mean making the same mistake over and over again[136].

The success of medical education is not measured on the ability of each individual to make improvement on their knowledge and skills at their own pace and to differing standards, but to gain competence and mastery in a tight time frame whilst providing a safe service.

The subject numbers have been small, but in a study by Alvand[137] 20 medical students were given two arthroscopic tasks (one shoulder and one knee designed to represent core skills required for arthroscopic training). Seven students were unable to achieve competence in the shoulder task and four in the knee task. Those students who achieved task competence also had better objective technical dexterity. The conclusion was that some individuals were unable to achieve competence despite sustained practice.

To test the impact of aptitude on the learning curve of laparoscopic suturing Buckley et al[138] from the National Training Centre at the Royal College of Surgeons in Ireland had recruited two group of ten medical students from opposite ends of the aptitude spectrum (visual-spatial ability, depth perception, and psychomotor ability). The Visuo-spatial aptitude was tested on a previously validated paper based tests selected from the kit for factor-referenced cognitive test. Psychomotor aptitude was tested by measurement of manual and finger dexterity and hand-eye coordination using a grooved pegboard (previously validated and well-recognised assessment tool). Spatial Aptitude was tested using a pictorial surface orientation test (PicSOr[37]). All medical students from both spectrums were tested consecutively using the ProMIS III laparoscopic simulator[139] that generates instrument tracking and provides measurements of performance. The results showed a large gap between the two groups. All in the high aptitude group achieved proficiency after a mean of 7 attempts (range, 4-10). In the low aptitude group 30% achieved proficiency after a mean of 14 attempts (range, 10-16), 40% demonstrated improvement but did not attain proficiency, and 30% failed to progress. The authors concluded that high aptitude is directly related to earlier completion of the learning curve and a proportion in the low aptitude group could not reach proficiency despite practice. This same author on a separate paper had compared the aptitude tests of medical students to those in higher surgical training[140], and found that there was a degree of self-selection of candidates with high innate ability into surgical training. The Higher Surgical Trainees scored higher (57.7% vs 31.1%) in the visuo-spatial aptitudes compared to surgical novices (medical students) but there were 11% of Higher Surgical Trainees who had scored lower than the average novices. The authors argue that to avoid candidates with low innate abilities to struggle after entering surgical training, testing during medical school and prior to specialisation may reduce future failure to successfully complete surgical training.

5.6 Limitations of the study design

The recruiting of participants (the sampling procedure) was non-random (nonprobability) as only those participants who self-referred were included. This convenience sampling method may contain sources of bias but is a way of recruitment that is commonplace in medical educational research as random (probability) sampling can be time-consuming and expensive, and not feasible in some situations. The OASC at RAF Cranwell introduces new tests as they are developed into the Aptitude Test without the candidate being able to differentiate which of the tests are counted towards the final mark. The results of the new tests are kept for longitudinal correlation and only when validated will be introduced into the pool of 'live' questions.

The sample size in the Basic Surgical Skills Simulation arm was much smaller than in the Flying Aptitude Test. The difficulty encountered in recruitment, time, organisation, location, large human resources and high cost is exactly why the development of a computerised test to assess technical potential for surgery as it is being used in the RAF will be of benefit.

Error data was collected in the laparoscopic simulated tests but the numbers were so small (only 2 out of the 177 participants in the Common Bile Duct Cannulation group had made a mistake) that statistical analysis were not carried out.

5.7 Future Directions

A future continuation prospective study for which ethical approval has already been granted and participant consent obtained in advance, will be able to look into a prospective study by following up the undergraduates into specialty training with the hypothesis of those performing well in the simulated tests will consequently score well in the intra-operative Procedure Based Assessments, linking the bench model into the clinical setting.

There was a marked difference in the self-rating results between the genders, with female participants reporting an unfounded lower expectation of their own performance. This study has shown an equally good performance from female medical students compared to their male peers in the Flying Aptitude Test as well as the Laparoscopic and Basic Surgical Skills Tests in this study, with significant correlation between the tests.

There is no evidence yet to use the Flying Aptitude Test for selection into surgical training, but once validated, a surgical aptitude test could be potentially incorporated into early post graduate training to encourage graduates into a career in surgery. Such test could support self-actualisation and empowerment of female trainees into believing in their own potential technical ability.

Appendix

Appendix 1: Basic Surgical Skills Assessment Sheet RCS Edinburgh

RCS Edinburgh Basic Surgical Skills Assessment Sheet – Day One

Date………… Trainee…………………………………………………..

	Score					Comments	Signature of trainer
Core Skills							
Hand tied reef knot/Surgeon's knot							
Correct knot tied							
Correct tension applied							
Instrument tie							
Correct knot tied							
Correct tension applied							
Suturing (interrupted/subcuticular)							
Correct bite placement							
Good tissue approximation							
Minor Surgery							
Excision of skin lesion/LA techniques/FNA/Trucut							
Handling instruments							
Handling tissues							
Wound Management							
Debridement traumatic wound/abscess drainage/debriding necrotic wound							
Handling instruments							
Handling tissues							
Diathermy							
Power settings							
Surface area							
Instrument placement							
Tissue conductivity							

Assessment of skill score 1. Inadequate, requires constant supervision 2. Adequate but needs further practice 3. Satisfactory

Appendix 2: Study Protocol Version 3

Study Protocol

A Pilot Study: The Use Of The RAF Flying Aptitude Test For The Selection Of Surgical Trainees

Chief Investigator:

Mr Charles Maxwell-Armstrong

Principal Investigador:

Ms Hyunmi Park

Surgical Directorate, Queens Medical Centre, Nottingham

Study sponsored by

Nottingham University Hospitals NHS Trust

REC Ref 10/H0401/43

NHS R&D Ref: 10GS002

Version 3: 10th Dec 2010

INDEX

Study Protocol

1. TITLE OF STUDY

The use of RAF Flying Aptitude Test for the selection of Surgical Trainees

2. NATURE OF PROJECT

Study using mixed quantitative/qualitative methodology

3. INVESTIGATORS

3a. Principal Investigator

Ms Hyunmi Park

Surgical Teaching & Research Fellow

Department of Surgery

E Floor, West Block, QMC

Nottingham University Hospitals

Nottingham NG7 2UH

Tel 01158 231172, Mob 07909 615687

hyunmi.park@nhs.net

3b. Chief Investigator & Supervisor

Mr Charles A Maxwell-Armstrong

Consultant colorectal surgeon

Department of Surgery

E Floor, West Block, QMC

Nottingham University Hospitals NHS Trust

Nottingham, NG7 2UH

Tel 01159249924

Charles.maxwell-armstrong@nuh.nhs.uk

3c. Supervisors

Prof Bryn Baxendale

Centre Director

Trent Simulation & Clinical Skills Centre

Nottingham University Hospital

Nottingham, NG7 2UH

01159 249924

Bryn.baxendale@nuh.nhs.uk

Prof. E Ferguson

Faculty of Psychology

University of Nottingham

Nottingham, NG7 2RD

01159 515327

eamonn.Ferguson@nottingham.ac.uk

Air Cdre W J Coker

OC RAF Centre of Aviation Medicine

The Library, Building 128

RAF Centre of Aviation Medicine

RAF Henlow

Bedfordshire, SG16 6DN

01462 851515 Ext 8020

cokerw567@henlow.raf.mod.uk

Dr R Dennick

Faculty of Medicine and Health Sciences

Medical Education Unit

University of Nottingham

Nottingham, NG7 2RD

01158 230013

reg.dennick@nottingham.ac.uk

4. PREFERRED TIMETABLE

4a. Preferred start date: 01/06/2010

4b. Expected date of project's completion: 31/12/2012

5. SPONSOR / OTHER ORGANISATIONS INVOLVED AND FUNDING

5a. Department/Organisation requesting research:

Surgical Directorate, Nottingham University Hospitals NHS Trust

Sponsored by Research and Development Department, Nottingham University Hospitals NHS Trust

PROTOCOL SIGNATURES

Chief and Principal Investigator Declaration:

I am aware of my responsibilities as Chief and Principal Investigator under the guidelines of Good Clinical Practice (GCP)[1], The Medicines for Human Use (Clinical Trials) Regulations 2004, SI 2004 No.1031, and the study protocol.

I agree to conduct the study in accordance with the aforementioned.

Chief Investigator:

Name: Mr Charles Maxwell-Armstrong **Title:** ______________________

Signature: ____________________ **Date:** _____________________

Principal Investigator:

Name: Ms Hyunmi Park **Title:** ________________________

Signature: _______________________ **Date:** ______________________

[1] ICH Harmonised Tripartite Guideline E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) Step 5, adopted by CPMP July 1996.

6. OTHER REC APPROVAL

Has the proposed study been submitted to any other reviewing body? If so, please provide details:
MoD Research Ethics Committee approval confirmed 7th Jan 2010 Ref:1002/299
Dr Robert Linton
MoD Research Ethics Committee (General)
MINISTRY OF DEFENCE
Level 1, Zone K, Main Building,
Whitehall
London SW1A 2HB

7. PURPOSE OF THE STUDY

To test and develop a valid and reliable aptitude test that is able to identify those subjects who can perform well in surgery especially in laparoscopic and endoscopic procedures

8. STUDY DESIGN, METHODOLOGY AND DATA ANALYSIS

Study Summary

There has been a revolutionary advance in surgery with the introduction of minimal access surgery (MAS) and use of technology in laparoscopic surgery. In more recent times the training time of surgeons has shortened due to the introduction of the European Working Time Directive (EWTD) and new run-through curriculum. The current method of surgical trainee selection emphasises on academic achievement, publications and research, which does not include any objective measure of their intra-operative performance or their innate skills.

The use of psychometric and aptitude testing is widely used in both military and industry personnel selection. All people applying to join the Royal Air Force or Royal Navy as aircrew are assessed using aptitude tests, which have been designed to determine the innate ability in areas that cannot be trained [(1)]. Our recent study has shown a statistically significant difference in simulated laparoscopic performance between Harrier pilots (selected through the RAF aptitude test) and undergraduate medical students [(2)].

The assessment methods within surgery are also evolving. The use of structured procedure based assessment (PBA) is expected from all junior trainees. This should help in providing a universal and objective measurement of their intra-operative skills.

Our study will measure the performance of participants of different surgical experience in the RAF Flying Aptitude Test and in simulated laparoscopic and endoscopic tests. For the surgical core trainee group their procedure based assessment results will be collected prospectively and compared to their scores of the simulated tests. Validated psychometric questionnaires will also be used.

We aim to identify a valid test capable of discriminating candidates of different surgical ability.

Background

The selection process of surgical trainees should identify subjects who ultimately become the best specialty consultants or consultant surgeons. The importance of selecting the right candidate is very important, as surgical training is lengthy and expensive. Choosing the wrong candidate can have a detrimental effect on patients, colleagues, hospital, the individual and their families. The cost of getting it wrong is very high [(3)]. The selection of surgical trainees within the National Health Service does not possess an assessment where to predict the expected surgical performance of trainees, although some have been suggested [(4,5)].

There has been a revolutionary advance in surgery over the last two decades with the introduction of minimal access surgery (MAS), which requires a new set of skills. Open surgery skills are not found to be transferable to MAS [(6)] and the traditional Halsteadian method of teaching open surgery did not seem to be suitable for learning MAS skills [(7)].

MAS is known to have great advantages in short-term patient outcomes and to be as oncologically safe as open surgery in cancer surgery. A drawback is that complication rates can be relatively high during the acquisition of MAS skills [(8)].

The training time of surgeons has shortened due to the introduction of the European Working Times Directive and the new run-through curriculum. The new structured curriculum selects candidates into surgical specialty (Core Training year 1) after only two years postgraduate experience (Foundation years). During the two foundation years, the exposure to surgery can be as short as four months. Having an interest in a career in surgery at this stage does not correlate with superior technical skills than those who chose other specialities [(9)].

The current selection checklist includes academic achievement, publications and research, which are easily measured parameters. There are no standardisation and no objective evaluation. This process may differ between regions or differ between years in the same region. This subjective, non-quantitative approach to selection and training in surgery may no longer be acceptable [(3)].

There have been various assessment tools in the form of box simulators and virtual simulators in laparoscopic surgery that have been validated in the training of MAS [(10-14)]. The performance in the simulated environment has been shown to correlate with real operating room performance and therefore the tests seem to be a valid way of assessing ability in a safe environment.

A Meta-analysis of 16 studies on the training effectiveness of virtual reality surgical simulators has found that training in VR simulators lessens the time and errors in the performance of a given surgical task and can clearly differentiate between less experienced and experienced trainees [15].

In both military and commercial aviation, the pilots are selected through an aptitude test which involves the assessment of their innate ability. The psychomotor assessments include the identification of those candidates with strong hand-eye coordination and spatial awareness [16-18].

This study will use the six part computer based FLYING APTITUDE TEST which includes the following components: Verbal Reasoning, Numerical Reasoning, Spatial Reasoning, Attentional Capability, Short Term Memory and Psychomotor Ability. The computer based test will last less than one hour and the instructions will be given prior to each test on screen.

Fundamental abilities such as psychomotor skills, visuospatial ability and depth perception are critically important for MAS, which include catheter-based interventions (e.g. vascular surgery), NOTES (Natural Orifice Transluminal Endoscopic Surgery) and robotic surgery [19]. Surgical leaders from around the world identified the following criteria as very important in the selection of surgical trainees: innate dexterity, spatial perception, hand-eye coordination, aiming, multi-limb coordination, hand-arm steadiness and ability to interpret and manipulate from images [20-21]. At the National Surgical Training Centre (Royal College of Surgeons Ireland, Dublin), Higher Surgical Trainee candidates undergo, as part of their selection process, objective assessment of their surgical technical skills in simulated environments [3] and psychomotor testing which include: the use of Pictorial Surface Orientation Test (PicSOr) [22] for perceptual ability and paper based tests for visuospatial ability [19].

As in the selection of pilots, future surgeons should be selected for their above-mentioned abilities. This will minimise selection of candidates who can struggle during surgical training and thereafter in surgical practice.

Project Objectives

Hypothesis

- The Flying Aptitude Test will be able to identify those participants with good spatial reasoning
- Those passing the Flying Aptitude Test will perform better at the simulated laparoscopic and endoscopic tests
- Those performing well in the simulated tests will consequently score well in the intra-operative Procedure Based Assessments

Methods

There are various components to this study. These can take place on different days and in different order depending on the availability of the participants. Each participant will only need to undertake each part of the study once.

1. All participants will undergo computer based RAF Flying Aptitude Test. This will take place at RAF Cranwell. Transport will be provided to and from Nottingham University Medical School. It will normally take place on a Friday morning and the participant should be back in Nottingham by lunchtime.

   Prior to the test each participant will be asked to rate how good they might be at the test on the 'Self Rating Document' and this result later will be compared to their actual score.

   After the test they will be asked to fill in an anonymous feedback form on the research day.

2. All participants will undergo four validated simulated laparoscopic tasks on a box trainer (with camera): Instruction sheets will be provided on the day. The four tasks are:

- Bean drop task
- Block move task
- Bile duct cannulation task
- Appendicectomy task

3. Trainees from any speciality may undergo simulated laparoscopic cholecystectomy or colectomy (VR simulator)
4. Trainees from any speciality may undergo simulated Upper gastro intestinal or Lower gastro intestinal endoscopy (VR simulator)
5. All volunteers will be asked to fill the following psychometric questionnaires:

- University of Wales Institute of Science and Technology (UWIST)-Mood Adjective Checklist (MACL)
- Situational Triggers of Aggressive Responses (STAR) questionnaire
- Fear of failure
- Absorption scale
- Goldberg big 5
- Previous experience of laparoscopic surgery and computer games

The laparoscopic box trainer, VR simulated laparoscopic and endoscopic tests will take place at the Trent Simulation Centre, which is located in the Queens Medical Centre Campus Postgraduate Centre. Each test will take around 30 minutes and the instructions will be given just before the test is initiated. The psychometric questionnaires will be filled before and after the simulated tests are carried out. In total the participants will be at the Simulation Centre for less than 2 hours.

6. Correlate their simulated tests performance with already existing Procedure Based Assessment (ISCP). This assessment is carried out by the trainee's individual assigned educational supervisor or clinical supervisor responsible for their education. With their consent, every time the participant completes a PBA, this will be forwarded via secure email (with password) to the Principal Investigator's NHS email to be added to the study's database.

Participant Safety

The participants will be monitored and supervised by the principal investigator or a member of the research team at all stages. This includes the computer based Flying Aptitude Test, the laparoscopic box trainer and VR endoscopic and laparoscopic tests. During the box trainer test there will be no handling of organic matter therefore no risk of infection will be present.

ABBREVIATIONS USED: **RAF**, Royal Air Force**, MAS**; minimal access surgery; **LS**, laparoscopic surgery; **VR**, virtual reality; **Sim**, simulator; **ISCP**, intercollegiate surgical curriculum project; **PBA**, procedure based assessment; **CT**, core training; **EWTD**, European working time directive

References

1. Pilot Training: Aircrew Training: Fleet Air Arm: Operations and Support: Royal Navy. Royal Navy website 2008 (accessed 6 Nov 2008) http://www.royal-navy.mod.uk/server/show/conWebDoc.2432
2. Submission 11268, Laparoscopic simulator performance: Harrier fighter pilots versus medical and non-medical undergraduates. Accepted for publication in the Bulletin of the Annals of the Royal College of Surgeons of England.
3. Gallagher AG, Neary P, Gillen P, Lane, B, Whelan A, Tanner W, Traynor O. Novel Method for Assessment and selection of trainees for higher surgical training in general surgery. ANZ J Surg 2008:78:282-290
4. Elneel FHF, Carter F, Tang B, Cuschieri A. Extent of innate dexterity and ambidexterity across handedness and gender: Implications for training in laparoscopic surgery. Surgical Endoscopy 2008; 22(1): 31-37

5. Francis N, Hanna G, Cresswell A, Carter F, Cuschieri A. The performance of master surgeons on standard aptitude testing. American Journal of Surgery 2001; 182: 30-33
6. Figert P, Park A, Witzke D, Schwartz R. Transfer of training in acquiring laparoscopic skills. J Am Coll Surg 2001; 193 (5): 533-537
7. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. Surg Endosc (2003); 17: 1525-1529
8. Southern Surgeons Club. The learning curve for laparoscopic cholecystectomy. Am J Surg 1995; 170: 55-59
9. Cope DH, Fenton-Lee D. Assessment of laparoscopic psychomotor skills in interns using the MIST virtual reality simulator: a prerequisite for those considering surgical training? ANZ J surg 2007: 78:291-296
10. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am J Surg (2006) 191: 128-133
11. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. BJS (2004); 91: 146-150
12. Aggarwal R, Black SA Hance JR, Darzi A, Cheshire NJW. Virtual reality simulation training can improve inexperience surgeons' endovascular skills. Eur J vasc Endovasc Surg (2006);l 31: 588-593
13. Hance J, Aggarwal R, Moorthy K, Munz Y, Undre S, Darzi A. Assessment of psychomotor skills acquisition during laparoscopic cholecystectomy courses. Am J Surg (2005); 190: 507-511
14. Seymour NE, Gallagher AG, Roman SA, O-Brien MK, Bansal VK, Andersen DK, Satava RM. Virtual reality training improves operating room performance. Annals of Surg (2002); 236 (4): 458-464
15. Haque S, Srinivasan S. A Meta-Analysis of the training Effectiveness of Virtual Reality Surgical Simulators. IEE Transactions of information technology in Biomedicine 2006; 10 (1): 51-58
16. Carretta TR, Ree MJ. US Air Force pilot selection tests: what is measured and is it predictive? Aviat Space Environ Med 1996; 67 (3): 279-283.
17. Royal Navy. Aircrew Officer (Pilots and Observer) : Careers / Roles : Royal Navy [document on the Internet]. Royal Navy online; 2009 [cited 2009 Jan 28]. Available from: http://www.royalnavy.mod.uk/server?show=nav.6056&tab=3.
18. Royal Navy. Selection Process For The Royal Navy and Royal Marines. 2009
19. Gallagher A, Leonard G, Traynor O. Role and feasibility of psychomotor and dexterity testing in selection for surgical training. ANZ J Surg (2009); 79: 108-113

20. Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? Am J Surg (2001); 182:110-16
21. Dalkey NC. The Delphi Method: An experimental study of group opinion. The RAND Corporation, 1969. RM-58888-PR
22. Gallagher AG, Cowie R, Crothers I, Jordan-Black JA, Satava RM. PicSOr An objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. Surg Endosc (2003); 17: 1468-1471

9. ETHICAL CONSIDERATIONS

MoD Ethical approval granted ref. 1002/299 (07/01/2010)

NHS Ethical approval granted ref. 10/H0401/43 (17/06/2010)

NUH sponsorship granted ref. 10GS002 (22/09/2010)

10. PARTICIPANTS TO BE STUDIED

Number of participants: 450

Lower age limit: 18

Upper age limit: 65

Gender: male & female

Please provide justification for the sample size:

Our previous published study compared Nottingham University Medical Undergraduates to UK RAF and Naval Harrier pilots on their simulated laparoscopic ability using the box trainer, which we will be using in this study. The numbers of participants were 25 in the medical student group and 9 in the Harrier pilot group. The results showed a statistical significant difference with p = <0.0101 using Kruskal-Wallis tests and UNIANOVA.

The sample sizes have been chosen as they are the probable trainee numbers available in the East Midlands deanery and undergraduate medical students.

The sample size should be sufficient to reveal statistical significance.

11. SELECTION CRITERIA

Participant Selection

In order to address the study outcomes, several distinct group of volunteers will be required:

- Medical students (n=200)
- Surgical and Medical Core Training (CT)year 1 and 2 (n=100)
- Specialty Registrars in years 3 to 8 in the following specialities that have the technical

abilities to undertake minimally access surgery or investigations: Surgery (laparoscopic surgery, endoscopy, cystoscopy, arthroscopy, bronchoscopy, interventional radiology), obstetric & gynaecology (laparoscopic surgery), respiratory medicine (bronchoscopy), gastroenterology (endoscopy) and radiology (interventional radiology) (n=100)

- RAF doctors in training (n=50)

1. Selection of Nottingham Medical student volunteers:

Age restriction 18-65

Registered undergraduate medical student studying Medicine at the University of Nottingham.

No previous laparoscopic/endoscopic experience.

Some medical experience and knowledge.

2. Selection of East Midlands deanery CT1 & CT2 surgical and medical trainees volunteers:

Age restriction 18-65

Registered medical practitioners with the GMC

Very limited previous laparoscopic/endoscopic experience

3. Selection of East Midlands deanery Specialty Registrars volunteers:

Age restriction 18-65

Registered medical practitioners with the GMC

Holder of a National Training Number

Experience of varying degrees in laparoscopy/endoscopy/bronchoscopy/interventional radiology

Trainees not yet in the specialist register

4. Selection of RAF doctors

Age restriction 18-65

Registered medical practitioners with the GMC

Trainees not yet in the specialist register

12. RECRUITMENT

12a. Describe how potential participants will be identified:

Group 1: Undergraduate register at University of Nottingham Medical school. Recruitment via email, poster and lecture presentation

Group 2: Core surgical and medical trainees identified via East Midlands Deanery with their permission. Recruitment via email and lecture presentation on training days

Group 3: Higher surgical, medical, O&G and radiology trainees identified via East Midlands Deanery with their permission. Recruitment via email and lecture presentation on training days

Group 4: RAF doctors in training. Recruited via Air Cdre Coker, OC RAF Centre of Aviation Medicine, RAF Henlow

12b. Describe how potential participants will be approached:

1. Advertising via individual undergraduate's, trainees' and RAF doctors' email

2. Posters at the University of Nottingham and teaching hospitals in the East Midlands Deanery.

3. Lecture presentation on undergraduate teaching day, doctors training day and RAF doctors meeting day

12c. Describe how potential participants will be recruited:

Email or phone reply with details

For the surgical simulation part of the study at the Trent Simulation Centre participants who have already been recruited via email replies will be sent a link to the Trent Simulation Centre research website which will host an online booking form.

The online booking form will involve the participant entering their contact details and choosing which date and time slot is suitable for them to attend.

There will be two links:

- A live link for Nottingham University Hospital staff for an online booking form hosted by the Trent Simulation Centre Research site. The slots available will be up to date and confirmation will be instant.
- The other link will be an online form which is accessible to anyone with internet connection but will ask the participant to enter two suitable dates and time slots which will then need approving by the principal investigator and confirmed via email.
- Both online forms will forward the details of the participants to the Principal investigator (Hyunmi Park on Hyunmi.park@nhs.net) and to the Clinical Skills Technician Julie Prince (Julie.prince@nuh.nhs.uk). All participants details will then be erased and not recorded in the website domain for security.

13. CONSENT

13a. Please describe the process you will use when seeking and obtaining consent:

The participant information sheet will be emailed at least 24 hours prior to the first test date. The participant will have informed written consent taken on the morning of the first study using the standard consent form. The participant will be notified that their name, date of birth and aptitude test scores will be kept by the Officers and Aircrew Selection Centre in RAF Cranwell. Participants who later wish to join the RAF and have to undertake an aptitude test will need to disclose of their previous experience in our study.

13b. Will the participants be from any of the following groups?

Under 18: No　　Prisoners: No　　Mental Illness: No

Learning disabilities: No

13c. Are there any special pressures that might make it difficult for people to refuse to take part in the study? How will you address such issues?

There are not special pressures known

14. PARTICIPANT INVOLVEMENT: RISKS, REQUIREMENTS AND BENEFITS

14a. What are the potential hazards, risks or adverse effects associated with the study?

No potential hazards have been identified

14b. Does your study involve invasive procedures such as blood taking, muscle biopsy or the administration of a medicinal product?

No

14c. Please name the locations or sites where the work will be done:

1. Officers and Aircrew Selection Centre, RAF College Cranwell, Sleaford, Lincolnshire, NG34 8GZ

2. Trent Simulation Centre, Queens Medical Centre Campus, Nottingham University Hospitals NHS Trust, Nottingham, NG7 2UH

14d. Will group or individual interviews / questionnaires discuss any topics or issues that might be sensitive, embarrassing or upsetting?

No

14e. Is it possible that criminal or other disclosures requiring action (e.g. evidence of professional misconduct) could take place during the study?

No

14f. Please describe any expected benefits to the research participant:

For the undergraduate and junior trainees (CT1 and CT2) this study might be their first encounter with minimally invasive surgical simulators. This can provide them with a taster of what their future training might involve.

For the more experienced trainees it will be an opportunity to undertake other types of simulators than the ones they are accustomed to.

14g. Under what circumstances might a participant not continue with the study or the study be terminated in part or as a whole?

Only if the participant wishes to terminate the study early

15. FINANCIAL INCENTIVES, EXPENSES AND COMPENSATION

15a. Will travelling expenses be given?

The transport to and from RAF Cranwell (Officers and Aircrew Selection Centre) will be provided.

15b. Is any financial or other reward, apart from travelling expenses, to be given to participants? If yes, please give details and justification:

No

15c. If this is a study in collaboration with a pharmaceutical company or an equipment or medical device manufacturer, please give the name of the company:

No

16. CONFIDENTIALITY, ANONYMITY AND DATA STORAGE

16a. What steps will be taken to ensure confidentiality (including the confidentiality and physical security of the research data)? Give details of the anonymisation procedures to be used, and at what stage they will be introduced:

Data management

The candidates' data will be entered by Ms Park (principal investigator) onto a password protected Microsoft Access database held on the Nottingham University Hospitals NHS Server. Only members of the research team will have access to the data. A separate candidate log will be held identifying the candidate's details together with their unique study number. The tests results and personal data generated will be entered into a database in which each candidate will be identified by number only. Once the study finishes, only the participant's number will be stored and personal details will be erased. The consent forms will be kept with the rest of the data (minus the personal data) for 10 years at the Surgical Department Research office in a locked cabinet. There will be an option on the consent form which the participants can choose to tick if they wish their details to remain for future longitudinal studies.

16b. Who will have access to the records and resulting data?

The chief investigator and supervisors. For the aptitude test results only, the Officers and Aircrew Selection Centre at RAF Cranwell.

16c. Where, and for how long, do you intend to store the consent forms and other records?

The NHS will store all records for 10 years in accordance with NHS ethical guidelines and the MoD will retain them for 100 years.

Appendix 3: Recruitment Poster

Nottingham University Hospitals NHS
NHS Trust

Appendix 4: Participant Information Sheet

Participant Information Sheet

A Pilot Study:
The Use Of The
RAF Flying Aptitude Test
For The Selection Of
Surgical Trainees

Chief Investigator:

Mr Charles Maxwell-Armstrong

Principal Investigator:

Ms Hyunmi Park

Surgical Directorate, Queens Medical Centre, Nottingham

Study sponsored by

Nottingham University Hospitals NHS
NHS Trust

Version 3. 8th Nov 2010

Ministry of Defence Research Ethics Committee Reference: 1002/299

NHS Research Ethics Committee Reference: 10/H0401/43

Version 3: 8th Nov 2010

Study Title: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees

Ministry of Defence Research Ethics Committee Reference: 1002/299

NHS Research Ethics Committee Reference: 10/H0401/43

INDEX

Version 3: 8th Nov 2010
Study Title: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees
Ministry of Defence Research Ethics Committee Reference: 1002/299
NHS Research Ethics Committee Reference: 10/H0401/43

Information for Participants

1. Study title

A Pilot Study: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees

2. Invitation to take part

We would like to invite you to take part in a research study. Before you decide, you need to understand why the research is being done and what it would involve for you. Please read the following information carefully. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you want to take part.

3. What is the purpose of the research?

The Aim of our study is to test and develop a valid and reliable aptitude test that is able to identify those subjects who can perform well in surgery especially in laparoscopic and endoscopic procedures

4. Who is doing this research?

The principal investigator of this research is:

Ms Hyunmi Park MBChB MRCSEng

Research & Teaching Surgical registrar

Surgical Directorate, West Block E floor

Queens Medical Centre

Nottingham University Hospitals NHS Trust

Derby Road, Nottingham, NG7 2UH, UK

Tel: 01158231172, 01159249924 switch to be put through to mobile

Email: hyunmi.Park@nhs.net

5. Why have I been invited to take part?

In this study we will compare your ability at different stages of your medical training. You have been contacted as you are an undergraduate at Nottingham University, a postgraduate trainee in the East Midlands Deanery or an RAF doctor in training.

6. Do I have to take part?

You should only participate if you want to; choosing not to take part will not disadvantage you in any way.

7. What will I be asked to do?

There are various components to this study. These can take place on different days and in different order depending on your availability. Each participant will only need to undertake each part of the study once.

1. You will undergo a computer based RAF Flying Aptitude Test. This is a computer based test with six components: Verbal Reasoning, Numerical Reasoning, Spatial Reasoning, Attentional Capability, Work Rate and Psychomotor Ability. It should last less than one hour and the instructions will be given prior to each test on screen. This will take place at RAF Cranwell. Transport will be provided to and from Nottingham University Medical School. It will normally take place on a Thursday morning and the participant should be back in Nottingham by lunchtime.

   Before the test you will be asked to rate yourself in how good you will do on a form titled 'Self rating document' which later we will compare to your actual score.

   After your test you will be asked to fill in an anonymous feedback form to rate various aspects of the research day

2. You will undergo four validated simulated laparoscopic tasks on a box trainer (with camera): Instruction sheets will be provided on the day. The four tasks are:

- Bean drop task
- Block move task
- Bile duct cannulation task
- Appendicectomy task

3. You may be asked to undertake a simulated laparoscopic cholecystectomy or colectomy (VR simulator)
4. You may be asked to undertake a simulated Upper gastro intestinal or Lower gastro intestinal endoscopy (VR simulator)

5. All participants will be asked to fill the following psychometric questionnaires. We wish to see if there is a correlation between any personality trait and the results of our simulated tests:

- University of Wales Institute of Science and Technology (UWIST)-Mood Adjective Checklist (MACL)
- Situational Triggers of Aggressive Responses (STAR) questionnaire
- Fear of failure
- Absorption scale
- Goldberg big 5
- You will be asked on your previous experience of laparoscopic surgery and computer games

The laparoscopic box trainer, VR simulated laparoscopic and endoscopic tests will take place at the Trent Simulation Centre, which is located in the Queens Medical Centre Campus Postgraduate Centre. Each test will take around 30 minutes and the instructions will be given just before the test is initiated. The psychometric questionnaires will be given before and after the simulated tests are carried out. In total the participants will be at the Simulation Centre for less than 2 hours.

6. PROSPECTIVE LONGITUDINAL STUDY: We wish to correlate your simulated test performance with the already existing Procedure Based Assessment (ISCP). This assessment is carried out by your individual assigned educational supervisor or clinical supervisor responsible for your education. With your consent, every time you complete a PBA, this will be forwarded via secure email (with password) to the Principal Investigator's NHS email to be added to the study's database.

Individual results will not be available to the candidates but once the study has been accepted for publication the participants will be contacted via email with the publication details

8. What is the device or procedure that is being tested?

Our study will measure the performance of individuals with different surgical experience in the RAF flying aptitude test, simulated laparoscopic and endoscopic tests. For the surgical core trainee group their procedure based assessment results will be collected prospectively and compared to their scores of the simulated tests. Validated psychometric questionnaires will also be used.

We aim to identify a valid test capable of discriminating candidates of different surgical ability

9. What are the benefits of taking part?

You will have an opportunity to experience the aptitude test that current UK military pilots have to undergo as part of their selection process.

For undergraduate and junior surgical and medical trainees, the laparoscopic and endoscopic simulators will give you an insight into your future training tools.

10. What are the possible disadvantages and risks of taking part?

The data collected from the computer based RAF Flying Aptitude Test will be held by the Officers and Aircrew Selection Centre and may affect your future score if you wish to join the UK military as a pilot in the future (this will not affect many of you as there is an age limit for recruitment for both the RAF and Navy).

11. Can I withdraw from the research and what will happen if I don't want to carry on?

You can withdraw at any point from the research and only the data collected up to that point will be included in the study. The personal details will be erased at the end of the study.

12. Are there any expenses and payments which I will get?

The transport to and from RAF Cranwell (Officers and Aircrew Selection Centre) may be provided.

13. Will my taking part or not taking part affect my medical career?

Taking part will have no effect on your medical career. The data collected from this study will not be available to any supervisor, deanery or university.

14. Whom do I contact if I have any questions or a complaint?

The first point of contact for all participants will be the Principal Investigator:

Ms H Park

Research & Teaching Surgical registrar

Department of Surgery,

E Floor, West Block,

Nottingham University Hospitals,

Nottingham, NG7 2UH

Email: hyunmi.Park@nhs.net

Tel: 01158231172, 01159249924 switch to be put through to mobile

If you have a complaint you can contact the following Independent researcher:

Mr M Robinson

Consultant colorectal surgeon

Department of Surgery

E Floor, West Block,

Queens Medical Centre,

Nottingham University Hospitals NHS Trust,

Nottingham, NG7 2UH

Email: mike.robinson@nuh.nhs.uk

Tel: 01159 249924

If you are a RAF doctor, you can also contact:

Air Cdre W J Coker

OC RAF Centre of Aviation Medicine

RAF Henlow

Email: cokerw567@henlow.raf.mod.uk

Tel: 01462 851515 Ext 8020

15. What happens if I suffer any harm?

This study does not involve any physical tests on yourselves but is a study testing your skills on different simulators. There is very low risk of harm, but if anything was to occur, please contact the Principal Investigator who will assist wherever possible.

For service personnel, If you suffer any harm as a result of taking part in this study, you can apply for compensation under the MoD's 'No-Fault Compensation Scheme'

16. Will my records be kept confidential?

All the data will be entered by Ms Park (principal investigator) onto a password protected Microsoft Access database held on the Nottingham University Hospitals NHS Server. Only members of the research team will have access to the data. A separate candidate log will be held identifying your details together with your unique study number. The tests results and personal data generated will be entered into a database in which each candidate will be identified by number only. This will be stored in a password protected Microsoft Access database held on the Nottingham University Hospitals NHS Server. Once the study finishes, only your identification number will be stored and personal details will be erased. The consent forms will be kept with the rest of the data (minus the personal data) for 10 years at the Surgical Department Research office in a locked cabinet. There will be an option on the consent form which you can choose to tick if you wish your details to remain for future longitudinal studies.

Your name, date of birth and results of the Flying Aptitude Test held at RAF Cranwell will be kept by the Officers and Aircrew Selection Centre. The reason for the RAF keeping the data is for future reference if you choose to join the British Armed forces as an aircrew, which will not be applicable to most of you as there is an age limit in recruitment.

17. Who is organising and funding the research?

The Nottingham University Hospitals NHS Trust and RAF Centre of Aviation Medicine

18. What will happen to the results of this study?

The study will be written up as a PhD/MSc thesis. The results will also be submitted for presentation in local, national and international meetings and for publication in professional medical or educational journals.

19. Who has reviewed the study?

This study has been approved by the MoD Research Ethics Committee with Ref 1002/299

The study has also been reviewed by the Derbyshire NHS Ethics Committee Ref 10/H0401/43

The study has also been reviewed by:

- Department of Surgery at the Nottingham University Hospitals NHS Trust
- Research and Development at the Nottingham University Hospitals NHS Trust

20. Further information and contact details

Ms H Park

Research & Teaching Surgical registrar

Department of Surgery, E Floor, West Block,

Nottingham University Hospitals, Nottingham, NG7 2UH

Email: hyunmi.Park@nhs.net

Tel: 0115 8231172, 01159249924 switch to be put through to mobile

21. Compliance with the Declaration of Helsinki

'This study complies and at all times will comply with the Declaration of Helsinki[1] as adopted at the 52nd WMA General Assembly, Edinburgh, October 2000 and with the Additional Protocol to the Convention on Human Rights and Biomedicine, concerning Biomedical Research, (Strasbourg 25.1.2005).

[1]World Medical Association (2000) Declaration of Helsinki. Ethical principles for medical research involving human subjects. 52nd World Medical Association General Assembly, Edinburgh, Scotland October 2000'.

Appendix 5: Consent for Participants

CONSENT FORM FOR PARTICIPANTS

Please initial the boxes below

☐ 1) The nature, aims and risks of the research have been explained to me. I have read and understood the Participant Information Sheet (v2 09.06.2010) and understand what is expected of me. All my questions have been answered fully to my satisfaction.

☐ 2) I understand that if I decide at any time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately without having to give a reason for my withdrawal.

☐ 3) I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

☐ 4) I agree to volunteer as a participant for the study described in the information sheet and I give full consent to my participation in this study.

☐ 5) This consent is specific to the particular experiment described in the Participant Information Sheet (v2 09.06.2010) attached and shall not be taken to imply my consent to participate in any subsequent experiment or deviation from that detailed here.

☐ 6) I understand that the results of this study will not be available to my trainers, university or deanery and cannot affect my medical training in the future.

☐ 7) I understand that data collected during the RAF selection test may be looked at and kept securely by individuals from the Officers & Aircrew Selection Centre and from the Ministry of Defence for up to 100 years. I give permission for these individuals to have access to the results of the study

☐ 8) I understand that the data collected during the whole study may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research in accordance to existing protocols. I give permission for these individuals to have access to the results of the study.

OPTIONAL

☐ 9) By ticking this box I agree for my personal details to be kept by the research team and to be contacted for longitudinal studies in the future (Re Information sheet page 3, part 6)

Participant's Statement: **I** ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Participant Information Sheet about the project, and understand what the research study involves.

Signed ____ Date _________________

Investigator's Statement: **I** ____

confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the Participant.

Signed ____ Date _________________

AUTHORISING SIGNATURES

The information supplied above is to the best of my knowledge and belief accurate. I clearly understand my obligations and the rights of research participants, particularly concerning recruitment of participants and obtaining valid consent.

Signature of Chief Investigator

Signed ____ Date _________________

Name and contact details of Principal Investigator:

Ms Hyunmi Park MBChB MRCSEng

Surgical Directorate

West Block E floor, Queens Medical Centre

Nottingham University Hospitals NHS Trust

Derby Road, Nottingham, NG7 2UH

Tel: 01159249924 Email: hyunmi.park@nhs.net

Appendix 6: Participant's questionnaire

Date

Study Title: Flying Aptitude Test for the selection of surgical trainees

Questions for Simulation Skills Participants

Name …………………… DoB……… Nationality ……………

Email …………………………………….

Mobile…………………….. Email 2 ………………………………….

- Are you: Male Female
- Interest in a surgical career Yes No
- Interest in a GI career Yes No
- Circle your speciality & year of training

Medical Student	1	2	3	4	5	6
Foundation	1	2				
Core Training						
Surgery	1	2				
Medicine	1	2				
Specialty training						
Surgery						
Cardiothoracic surgery	1	2	3	4	5	6
General surgery	1	2	3	4	5	6
Neurosurgery	1	2	3	4	5	6
Oral and maxillofacial surgery	1	2	3	4	5	6
Otolaryngology (ENT)	1	2	3	4	5	6
Paediatric surgery	1	2	3	4	5	6
Plastic surgery	1	2	3	4	5	6
Trauma and orthopaedic surgery	1	2	3	4	5	6
Urology	1	2	3	4	5	6
Academic surgery	1	2	3	4	5	6
Obstetric & Gynaecology	1	2	3	4	5	6
Medicine						
Respiratory	1	2	3	4	5	6
Gastro Intestinal	1	2	3	4	5	6
Radiology	1	2	3	4	5	6

- If you are in a speciality providing the following procedures, please indicate how many you have undertaken:

Laparoscopy surgery	1~10	11~25	26~50	51~75	76~100	>100
Endoscopy	1~10	11~25	26~50	51~75	76~100	>100
Arthroscopy	1~10	11~25	26~50	51~75	76~100	>100
Bronchoscopy	1~10	11~25	26~50	51~75	76~100	>100
Cystoscopy	1~10	11~25	26~50	51~75	76~100	>100
Interventional radiology	1~10	11~25	26~50	51~75	76~100	>100

Have you any experience of games played on computers or consoles?

Yes No

If yes, how would you rate your ability?

Beginner Average Intermediate Expert

What type of games do you play?

Role play War/Fighting games Racing Sports Arcade

Can you use chopsticks?

Yes No

If yes, are You a…..

Beginner Average Intermediate Expert

Appendix 7: Ministry of Defence Ethical Approval

MOD Research Ethics Committee (General)

Corporate Secretariat
Bldg 5, G01-614
Dstl Porton Down
Salisbury, Wiltshire
SP4 0JQ

Secretary: Marie Jones
telephone: 01980 658155
e-mail: mnjones@dstl.gov.uk

Ms H Park MBChB MRCSEng — Ref: 1002/299
Research & Teaching Surgical Registrar
Department of Surgery,
E Floor, West Block,
Nottingham University Hospitals,
Nottingham, NG7 2UH — 7TH January 2010

Dear Ms Park,

Re: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees – version 3 (ref:1002/299)

Thank you for submitting this interesting protocol for ethical review and for making minor revisions.

I am now happy to confirm ethical approval for this research and should be grateful if you would send me a copy of your final report on completion of the study.

This approval is conditional upon adherence to the protocol – please let me know if any amendment becomes necessary.

I note that you have also applied for approval from your NHS REC and that you will send me the recruitment poster and e-mail when they are available.

I hope the research goes well.

Yours sincerely,

Robert Linton

Dr Robert Linton

Chairman MoD Research Ethics Committee (General)

MoD Ethical approval for 1st Amendment

MOD Research Ethics Committee (General)

Corporate Secretariat
Bldg 5, G01-614
Dstl Porton Down
Salisbury, Wiltshire
SP4 0JQ

Secretary: Marie Jones
telephone: 01980 658155
e-mail: mnjones@dstl.gov.uk

Ms H Park MBChB MRCSEng Ref: 1002/299
Clinical Teaching Fellow
General Surgery
E floor, West Block
Queen's Medical Centre
Nottingham University Hospitals NHS Trust
Derby Road
Nottingham
NG7 2UH

Dear Ms Park,

Re: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees – version 3 (ref:1002/299) – 1st amendment

Thank you for your e-mail requesting ethical approval of some changes to this protocol. These include the use of two further forms, an online booking system and recruitment of medical as well as surgical trainees.

I agree that this will add value to the study and on behalf of MODREC am happy to give ethical approval for the amendment.

Yours sincerely,

Robert Linton

Dr Robert Linton

Chairman MOD Research Ethics Committee (General)

telephone: 020 8877 9329
e-mail: robert@foxlinton.org

mobile: 07764616756

MoD Ethical approval for 2nd amendment

MOD Research Ethics Committee (General)

Corporate Secretariat
Bldg 5, G01-614
Dstl Porton Down
Salisbury, Wiltshire
SP4 0JQ

Secretary: Marie Jones
telephone: 01980 658155
e-mail: mnjones@dstl.gov.uk

Ms H Park MBChB MRCSEng
Research & Teaching Surgical Registrar
Department of Surgery,
E Floor, West Block,
Queen's Medical Centre
Nottingham University Hospitals
Derby Road
Nottingham
NG7 2UH

Ref: 1002/299

20th September 2011

Dear Ms Park,

Re: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees – version 4 (ref:1002/299) – 2nd amendment

Thank you for your e-mail requesting a second amendment to this protocol. You wish to recruit more widely and to use the scores that your participants obtained in the Basic Surgical Skills course. You have updated the protocol and the Participant Information Sheet to reflect these changes.

I agree that this will add value to the study and on behalf of MODREC am happy to give ethical approval for this amendment.

Yours sincerely,

Robert Linton

Dr Robert Linton
Chairman MOD Research Ethics Committee (General)

telephone: 020 8877 9329
e-mail: robert@foxlinton.org
mobile: 07764616756

Appendix 8: NHS Ethical Approval

NHS
National Research Ethics Service

Derbyshire Research Ethics Committee
1 Standard Court
Park Row
Nottingham
NG1 6GN

Telephone: 0115 8839461
Facsimile: 0115 9123300

17 June 2010

Mr Charles Maxwell-Armstrong
Consultant colorectal surgeon
Nottingham University Hospitals NHS Trust
Queens Medical Centre
Derby Road
Nottingham
NG7 2UH

Dear Mr Maxwell-Armstrong

Study Title:	**The use of RAF Flying Aptitude Test for the selection of Surgical Trainees**
REC reference number:	**10/H0401/43**

Thank you for your letter of 14 June 2010, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. *Where the only involvement of the NHS organisation is as a Participant Identification*

Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
REC application	35194/117639/1/838	28 April 2010
Letter from Sponsor		29 March 2010
Questionnaire: Questions for Simulation Skills Participants	1	27 March 2010
Questionnaire: Absorption Scale	1	27 March 2010
Questionnaire: Fear of Failure	1	27 March 2010
Questionnaire: Goldberg BIG 5 Markers	1	27 March 2010
Questionnaire: Star	1	27 March 2010
Investigator CV		30 April 2010
Questionnaire: UWIST Mood Adjective Checklist	1	27 March 2010
CV - Hyun-Mi Carty		
Certificate - Hyun-Mi Carty		19 May 2008
Certificate - Mr Charles Maxwell-Armstrong		22 January 2009
Advertisement		
PowerPoint Presentation Slide View		
The Royal College of Surgeons of England Bulletin		
Submission to MoD Research Ethics Committee		07 January 2010
Protocol	2	09 June 2010
Participant Information Sheet	2	09 June 2010
Participant Consent Form	2	09 June 2010
Response to Request for Further Information		14 June 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

10/H0401/43:	**Please quote this number on all correspondence**

Yours sincerely

Mr Phil Hopkinson
Chair

E-mail: Sam.Tuite@nottspct.nhs.uk

Enclosures:	*List of names and professions of members who took part in the review*
Copy to:	*Dr Maria Koufali, Research & Development Nottingham University Hospitals NHS Trust*

NHS Ethical Approval for 1st Substantial Amendment

Derbyshire Research Ethics Committee
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Tel: 0115 8839436
Fax: 0115 9123300

17 December 2010

Mr Charles Maxwell-Armstrong
Consultant colorectal surgeon
Nottingham University Hospitals NHS Trust
Queens Medical Centre
Derby Road
Nottingham
NG7 2UH

Dear Mr Maxwell-Armstrong

Study title:	**The use of RAF Flying Aptitude Test for the selection of Surgical Trainees**
REC reference:	**10/H0401/43**
Protocol number:	**1**
Amendment number:	**1**
Amendment date:	**27 October 2010**

Thank you for submitting the above amendment, which was received on 17 December 2010. I can confirm that this is a valid notice of a substantial amendment and will be reviewed by the Sub-Committee of the REC at its next meeting.

Documents received

The documents to be reviewed are as follows:

Document	Version	Date
Feedback form	1	03 November 2010
How good do you think you will be	1	03 November 2010
On-Line Booking Form	1	27 October 2010
Questionnaire: For Simulation Skills Participants	3	10 December 2010
Participant Information Sheet	3	08 November 2010
Protocol	3	10 December 2010
Notice of Substantial Amendment (non-CTIMPs)		

Notification of the Committee's decision

The Committee will issue an ethical opinion on the amendment within a maximum of 35 days from the date of receipt.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval for the research.

10/H0401/43: **Please quote this number on all correspondence**

Yours sincerely

Mrs Carol Marten
Committee Co-ordinator

E-mail: carol.marten@nottspct.nhs.uk

2nd amendment NHS Ethical Approval

NHS
National Research Ethics Service
NRES Committee East Midlands - Derby 2
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Tel: 01158839437
Fax: 01159123300

20 September 2011

Mr Charles Maxwell-Armstrong
Consultant colorectal surgeon
Queens Medical Centre
Derby Road
Nottingham
NG7 2UH

Dear Mr Maxwell-Armstrong,

Study title:	**The use of RAF Flying Aptitude Test for the selection of Surgical Trainees**
REC reference:	**10/H0401/43**
Protocol number:	**1**
Amendment number:	**2**
Amendment date:	**08 September 2011**

The above amendment was reviewed at the meeting of the Sub-Committee held on 20 September 2011.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Participant Information Sheet	4	01 September 2011
Protocol	4	01 September 2011
Notice of Substantial Amendment (non-CTIMPs)	2	08 September 2011
Covering Letter		08 September 2011

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

This Research Ethics Committee is an advisory committee to the East Midlands Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

10/H0401/43	**Please quote this number on all correspondence**

Yours sincerely

pp L Gregory

Mr Phil Hopkinson
Chair

Email: lisa.gregory@nottspct.nhs.uk

Enclosures: "After ethical review – guidance for researchers" *SL- AR2*

Appendix 9: Self Rating Form

Date:

Nottingham University Hospitals NHS
NHS Trust

Study Title: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees

Name: _________________________

How good do you think you will be?

Very bad ____ Very good

Self Rating document. v1: 3rd Nov 2010

Appendix 10: Flying Aptitude Test Feedback Form

Date: __________

Study Title: The use of RAF Flying Aptitude Test for the selection of Surgical Trainees

FEEDBACK FORM

Grade:

1) The contents of the Participant Information Sheet was informative

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

2) The Consent form was easy to understand

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

OFFICERS & AIRCREW SELECTION CENTRE

3) Overall, how satisfied are you with the visit to RAF Cranwell's OASC?

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

4) How satisfied are you with the travel arrangements to and from RAF Cranwell?

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

5) How satisfied are you with the access (via security) to RAF Cranwell?

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

6) How satisfied are you with the instructions given to you prior to the test?

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

RELEVANCE

7) It was interesting to see the selection process of current Aircrew selection in the UK

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

8) I can see the relevance of this study in the future of Medical/Surgical Education

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

9) Selection into UK Surgical Training may benefit with the introduction of an aptitude test as an addition to the already existing Interview process

	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	

10) Have you enjoyed participating in this study?

Yes	No	Don't know	Don't want to answer

11) Would you recommend it to your colleagues?

Yes	No	Don't know	Don't want to answer

Any comments for the Principal Investigator

Any comments for RAF Cranwell

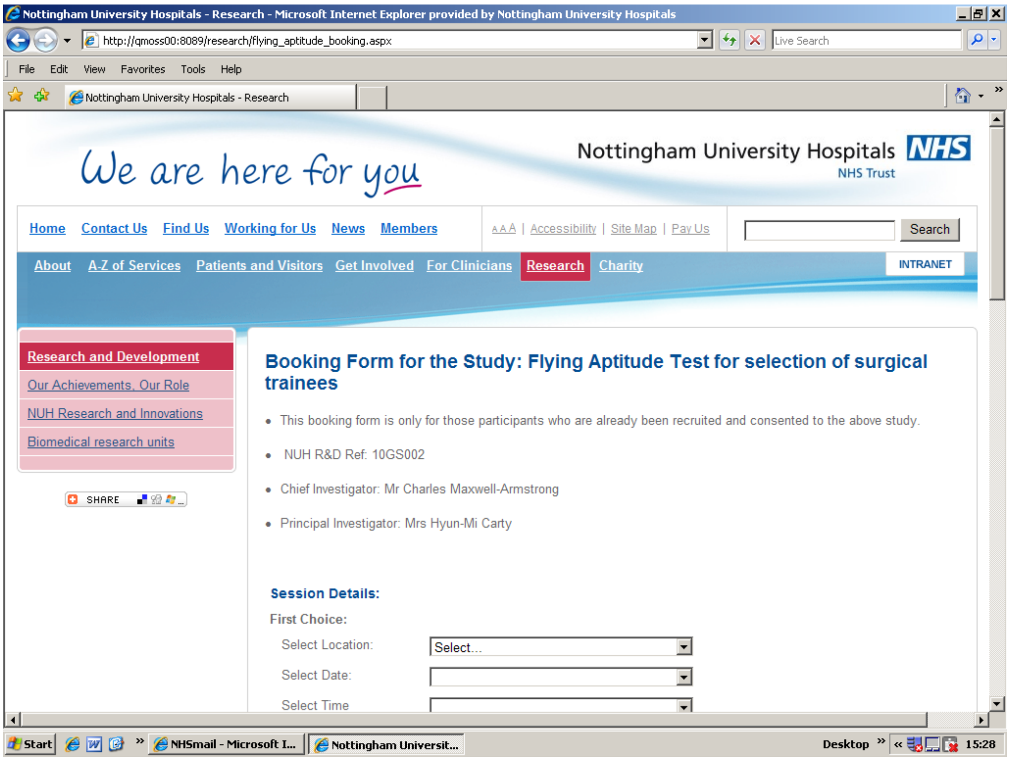
Any other comments

Thank you very much!

Ms Hyunmi Park

Principal Investigator hyunmi.park@nhs.net

Appendix 11: Online booking form for participants



Appendix 12: BSS results frequencies Female undergraduates

		Hand tie Correct	Hand tie tension	Instrument tie Correct	Instrument tie tension	Suturing tissue handling	Suturing Time & Motion	Suturing Instrument handling	Lesion excision Tissue handling	Lesion excision Time & Motion	Lesion Instrument handling	Tie score using BSS system (4-12)	suturing & lesion score using OASTS system (6-30)	Total BSS score (10-42)	Total BSS percent
N	Valid	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	Missing	98	98	98	98	98	98	98	98	98	98	98	98	98	98
Mean		2.56	2.44	2.78	2.83	4.00	3.72	3.72	4.28	4.39	4.39	10.61	24.50	35.11	78.4744
Median		3.00	3.00	3.00	3.00	4.00	4.00	3.00	5.00	5.00	5.00	11.00	26.00	37.00	84.3800
Mode		3	3	3	3	4	3[b]	3	5	5	5	11[b]	26[b]	38	87.50
Std. Deviation		.726	.726	.441	.354	.707	.905	1.034	.972	.993	.993	1.728	4.514	5.578	17.43059
Variance		.528	.528	.194	.125	.500	.819	1.069	.944	.986	.986	2.986	20.375	31.111	303.826
Range		2	2	1	1	2	3	3	3	3	3	6	14	18	56.25
Minimum		1	1	2	2	3	3	3	3	3	3	7	16	22	37.50
Maximum		3	3	3	3	5	5	5	5	5	5	12	29	40	93.75
Sum		23	22	25	26	36	34	34	39	40	40	96	221	316	706.27
Percentiles	25	2.00	2.00	2.50	2.75	3.50	3.00	3.00	3.50	3.50	3.50	10.00	21.50	33.00	71.8800
	50	3.00	3.00	3.00	3.00	4.00	4.00	3.00	5.00	5.00	5.00	11.00	26.00	37.00	84.3800
	75	3.00	3.00	3.00	3.00	4.50	4.50	5.00	5.00	5.00	5.00	12.00	28.00	38.50	89.0650

BSS results frequencies Male undergraduates

		Hand tie Correct	Hand tie tension	Instrument tie Correct	Instrument tie tension	Suturing tissue handling	Suturing Time & Motion	Suturing Instrument handling	Lesion excision Tissue handling	Lesion excision Time & Motion	Lesion Instrument handling	Tie score using BSS system (4-12)	suturing & lesion score using OASTS system (6-30)	Total BSS score (10-42)	Total BSS percent
N	Valid	13	13	13	13	13	13	13	13	13	13	13	13	13	13
	Missing	91	91	91	91	91	91	91	91	91	91	91	91	91	91
Mean		2.46	2.23	2.92	2.69	3.08	3.27	3.04	3.65	3.77	3.50	10.31	20.31	30.62	64.4238
Median		3.00	2.00	3.00	3.00	3.50	3.00	3.00	4.00	4.00	4.00	10.00	23.00	34.00	75.0000
Mode		3	2	3	3	4	3[b]	2	4	3	5	10[b]	14[b]	24[b]	43.75[b]
Std. Deviation		.776	.725	.277	.480	1.367	1.129	1.198	1.049	.992	1.541	1.932	6.527	8.234	25.73043
Variance		.603	.526	.077	.231	1.869	1.276	1.436	1.099	.984	2.375	3.731	42.606	67.798	662.055
Range		2	2	1	1	4	4	4	4	3	4	6	20	25	78.12
Minimum		1	1	2	2	1	1	1	1	2	1	6	8	15	15.63
Maximum		3	3	3	3	5	5	5	5	5	5	12	28	40	93.75
Sum		32	29	38	35	40	43	40	48	49	46	134	264	398	837.51
Percentiles	25	2.00	2.00	3.00	2.00	1.50	2.50	2.00	3.00	3.00	2.00	10.00	14.00	24.00	43.7500
	50	3.00	2.00	3.00	3.00	3.50	3.00	3.00	4.00	4.00	4.00	10.00	23.00	34.00	75.0000
	75	3.00	3.00	3.00	3.00	4.00	4.00	4.00	4.25	4.75	5.00	12.00	26.00	37.50	85.9400

List of Tables

List of Figures

Bibliography

1. Hall JC, Ellis C, Hamdorf J. Surgeons and cognitive processes. *Br J Surg* 2003;**90**:10–6.

2. Tsim P, Carty H-M, Baxendale B *et al*. Laparoscopic simulator performance: Harrier fighter pilots versus medical and non-medical undergraduates. *Bull R Coll Surg Engl* 2010;**92**:58–61.

3. Fielding D, Ferguson E, Baxendale B *et al*. Assessing laparoscopic skills in the novice. *Bull R Coll Surg Engl* 2008;**90**:246–9.

4. Galasko CS. Competence of pilots and surgeons. *BMJ Br Med J* 2000;**320**.

5. Innes D. Surgeons should have to demonstrate competence as pilots do. *BMJ* 2000;**320**:1145–6.

6. Galasko CS. Competencies required to be a competent surgeon. *Ann R Coll Surg Engl* 2000;**82**:89–90.

7. Beard JD, Jolly BC, Newble DI *et al*. Assessing the technical skills of surgical trainees. *Br J Surg* 2005;**92**:778–82.

8. GMC. Good Medical Practice (2013). 2013.

9. Bailey M, Cranwell RAF. Evolution of Aptitude Testing in the RAF. 1999:9–11.

10. Darley J. *The Reliability of the Tests in the Standard Battery. In Research Studies in Individual Diagnosis. Vol3, No4. DG Paterson (Ed).* Minneapolis: University of Minnesota Press, 1934.

11. Salvendy G, Stewart GK. The Prediction of Operator Performance on the Basis of Performance Tests and Biological Measures. *Proc Hum Factors Ergon Soc Annu Meet* 1975;**19**:457–66.

12. Adams JA. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychol Bull* 1987;**101**:41–74.

13. RAF. RAF recruitment eligibility check.

14. RAF. *RAF Notes for Guidance for OASC Aptitude Testing*., 2015.

15. Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: From selection to certification. *Curr Probl Surg* 2002;**39**:583–659.

16. Bell J. Royal Air Force Selection Procedures. *Ann R Coll Surg Engl* 1988;**70**:270–5.

17. Polavarapu H V, Kulaylat AN, Sun S *et al*. 100 years of surgical education: the past, present, and future. *Bull Am Coll Surg* 2013;**98**:22–7.

18. Fitzgerald JEF, Giddings CEB, Khera G *et al*. Improving the future of surgical training and education: Recommendations from the Association of Surgeons in Training. *Int J Surg* 2012;**10**:389–92.

19. Specialty Recruitment Applicant Handbook. 2014:0–57.

20. Surgeons RC of. *Improving Surgical Training. Proposal for a Pilot Surgical Training Programme.*, 2015.

21. Greenway D. *Shape of Training. Securing the Future of Excellent Patient Care.*, 2013.

22. Royal College of Surgeons. Pilot sites announced for new surgical training programme. 2017.

23. HEE. Person Specification General Surgery ST3 Career Progression. Health Education England. 2017.

24. Royal College of Surgeons. FRCS Fellowship of the Royal College of Surgeons. Intercollegiate Specialty Fellowship Examination (JCIE).

25. Bulstrode C, Hunt V. Selecting the best from the rest. *Surgeon* 2003;**1**:328–31.

26. Bloom BS, Engelhard MD, Furst EJ *et al*. Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I Cognitive Domain. *New York* 1956;**16**:207.

27. Atherton JS. Learning and Teaching. Bloom's Taxonomy of Learning. 2013.

28. McKimm J, Swanwick T. Setting learning objectives. *Br J Hosp Med (Lond)* 2009;**70**:406–9.

29. Guide to the intercollegiate MRCS. 2013.

30. Pannabecker JR. Diderot, the Mechanical Arts, and the Encyclopedie: In Search of the Heritage of Technology Education. *J Technol Educ* 1994;**6**, DOI: 10.21061/jte.v6i1.a.4.

31. Harden RM, Stevenson M, Downie WW *et al*. Assessment of clinical competence using objective structured examination. *Br Med J* 1975;**1**:447–51.

32. Gallagher AG, Neary P, Gillen P *et al*. Novel method for assessment and selection of trainees for higher surgical training in general surgery. *ANZ J Surg* 2008;**78**:282–90.

33. Tansley P, Kakar S, Withey S *et al*. Visuospatial and technical ability in the selection and assessment of higher surgical trainees in the London deanery. *Ann R Coll Surg Engl* 2007;**89**:591–5.

34. Reznick R, Regehr G, MacRae H *et al*. Testing technical skill via an innovative “bench station” examination. *Am J Surg* 1997;**173**:226–30.

35. Shaharan S, Neary P. Evaluation of surgical training in the era of simulation. *World J Gastrointest Endosc* 2014;**6**:436–47.

36. OSATS ST3 Interview Scoring Criteria for Technical Skills - Cardiothoracic Surgery. 2015.

37. Gallagher AG, Cowie R, Crothers I *et al*. PicSOr: An objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. *Surg Endosc Other Interv Tech* 2003;**17**:1468–71.

38. Reinhardt-Rutland A, Gallagher A. *Visual Depth Perception in Minimally Invasive Surgery. In SA Robertson (Ed). Contemporary Ergonomics*. London: Taylor & Francis, 1995.

39. Dawson S. Perspectives on performance assessment in medical simulation. *Surgeon* 2011;**9**:S21–2.

40. Sturm LP, Windsor JA, Cosman PH *et al*. A Systematic Review of Skills Transfer After Surgical Simulation Training. *Ann Surg* 2008;**248**:166–79.

41. Dawe SR, Windsor JA, Broeders JAJL *et al*. A Systematic Review of Surgical Skills Transfer After Simulation-Based Training. *Ann Surg* 2014;**259**:236–48.

42. Tsim P, Ferguson E, Carty H-M *et al*. Is Laparoscopic Simulator Performance Affected by Training Environment? *Bull R Coll Surg Engl* 2011;**93**:169–72.

43. Tang CG, Hilsinger RL, Cruz RM *et al*. Manual Dexterity Aptitude Testing. *JAMA Otolaryngol Neck Surg* 2014;**140**:243.

44. Leblanc F, Champagne BJ, Augestad KM *et al*. A comparison of human cadaver and augmented reality simulator models for straight laparoscopic colorectal skills acquisition training. *J Am Coll Surg* 2010;**211**:250–5.

45. Wyles SM, Miskovic D, Ni Z *et al*. Analysis of laboratory-based laparoscopic colorectal surgery workshops within the English National Training Programme. *Surg Endosc Other Interv Tech* 2011;**25**:1559–66.

46. Giger U, Frésard I, Häfliger A *et al*. Laparoscopic training on Thiel human cadavers: A model to teach advanced laparoscopic procedures. *Surg Endosc Other Interv Tech* 2008;**22**:901–6.

47. RCSEd. Basic Surgical Skills Course.

48. Bloom BS, Englehard MD, Furst EJ *et al*. Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I Cognitive Domain. *New York* 1956;**16**:207.

49. Zawisza K (University of A. Bloom's Taxonomy of Learning. University of Arkansas. 2017.

50. Zawisza K (University of A. Bloom's Taxonomy of Learning. University of Arkansas. 2017.

51. Simpson E. *The Classification of Educational Objectives in the Psychomotor Domain: The Psychomotor Domain. Vol. 3*. Washington DC: Gryphon House, 1972.

52. Harrow A. *A Taxonomy of the Psychomotor Domain. A Guide for Developing Behavioral Objectives*. New York: McKay, 1972.

53. Dave R. *Developing and Writing Behavioural Objectives (R J Armstrong, Ed.)*. Educational Innovators Press, 1975.

54. Fitts P, Posner M. *Human Performance*. Belmont, CA: Brooks/Cole, 1967.

55. Reznick RK, MacRae H. Teaching Surgical Skills — Changes in the Wind.

N Engl J Med 2006;**355**:2664–9.

56. Ackerman PL. Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *J Exp Psychol Gen* 1988;**117**:288–318.

57. Richman H, Gobet F, Staszewski J *et al*. The Road to excellence: The acquisition of expert performance in the arts and sciences, sports and games. *Perceptual and Memory Processes in the Acquisition of Expert Performance: The EPAM Model. In K.A. Ericson (Ed.)*. Mahwah, MJ: Erlbaum, 1996, 167–87.

58. Schneider W, Shiffrin RM. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol Rev* 1977;**84**:1–66.

59. Thorndike EL. *Animal Intelligence : An Experimental Study of the Associative Processes in Animals / by Edward L. Thorndike*. New York : Macmillan, 1898.

60. Ericsson KA, Charness N, Feltovich PJ *et al*. *The Cambridge Handbook of Expertise and Expert Performance*. Ericsson KA, Charness N, Feltovich PJ, et al. (eds.). Cambridge: Cambridge University Press, 2006.

61. Procedure Based Assessment (PBA) Intercollegiate Surgical Curriculum Programme. 2017.

62. Direct Observation of Procedural Skills (DOPS). Intercollegiate Surgical Curriculum Programme. 2017.

63. Darzi A, Datta V, Mackay S. The challenge of objective assessment of surgical skill. *Am J Surg* 2001;**181**:484–6.

64. Hamdorf JM, Hall JC. Acquiring surgical skills. *Br J Surg* 2000;**87**:28–37.

65. Melton A. *Army Air Forces Aviation Psychology Program Research Reports: Apparatus Tests (Rep. No. 4)*. Washington DC: U.S. Government Printing Office, 1947.

66. Ackerman PL, Cianciolo AT. Psychomotor abilities via touch panel testing: Measurement innovations, construct, and criterion validity. *Hum Perform* 1999;**12**:231–73.

67. Hines M, O-Connor J. A measure of finger dexterity. *J Res Pers* 1926;**4**:379–82.

68. Tiffin J, Greenly RJ. Employee selection tests for electrical fixture assemblers and radio assemblers. *J Appl Psychol* 1939;**23**:240–63.

69. Blum ML. A contribution to manual aptitude measurement in industry. *J Appl Psychol* 1940;**24**:381–416.

70. Hayes EG. Selecting women for shop work. *J Res Pers* 1932;**11**:69–85.

71. Rim Y. The predictive validity of seven manual dexterity tests. *Psychologia* 1962;**5**:52–5.

72. Spencer FC. Teaching and measuring surgical techniques, the technical evaluation of competence. *Bull Am Coll Surg* 1978;**63**:9–12.

73. Kirby TJ. Dexterity testing and residents' surgical performance. *Trans Am Ophthalmol Soc* 1979;**77**:294–307.

74. Anastasi A. *Psychological Testing 3rd Ed*. New York: Macmillan Company, 1968.

75. Roman F. On manual dexterity. *Br J Ophthalmol* 1993;**77**:656.

76. Gardner H. *The Mind's New Science*. New York: Basic Books, 1987.

77. Assael LA. Every Surgeon Needs Mentors: A Halsteadian/Socratic Model in the Modern Age. *J Oral Maxillofac Surg* 2010;**68**:1217–8.

78. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: Rigorous science for the assessment of surgical education and training. *Surg Endosc Other Interv Tech* 2003;**17**:1525–9.

79. Unit RP. Medical Education England Review of the Impact of the European Working Time Directive (EWTD) on the Quality of Postgraduate Training. *rcseng.ac.uk* 2010.

80. Lowry J, Cripps J. Results of the online EWTD trainee survey. *Bull R Coll Surg Engl* 2005;**87**:86–7.

81. Simpson C, Cottam H, Fitzgerald JE *et al*. The European working time directive has a negative impact on surgical training in the UK. *Surgeon* 2011;**9**:56–7.

82. House of Commons. House of Commons debate . Q10 European Working

Time Directive. 2012.

83. Fitzgerald JE., Marron CD, Giddings CEB. The influence of specialty, grade, gender and deanery on the implementation and outcomes of European working time regulations in surgery. *Br J Surg* **98**:21.

84. ASiT. EWTD for Surgical Trainees. 2009.

85. Tait MJ, Fellows G a, Pushpananthan S *et al*. Current neurosurgical trainees' perception of the European Working Time Directive and shift work. *Br J Neurosurg* 2008;**22**:28-31-3.

86. Martin IC, Mason DG, Stewart J *et al*. *Emergency Admissions: A Journey in the Right Direction? A Report of the National Confidential Enquiry into Patient Outcome and Death*., 2007.

87. Morrow G, Burford B, Carter M *et al*. The Impact of the Working Time Regulations on Medical Education and Training: Literature Review. A Report for the General Medical Council. 2012.

88. Kelly BD, Curtin PD, Corcoran M. The effects of the European working time directive on surgical training: The basic surgical trainee's perspective. *Ir J Med Sci* 2011;**180**:435–7.

89. Lonergan PE, Mulsow J, Tanner WA *et al*. Residents' Operative Experience in General Surgery Programs. *Ann Surg* 2010;**251**:182.

90. Benes V. The European Working Time Directive and the Effects on

Training of Surgical Specialists (Doctors in Training). *Acta Neurochir (Wien)* 2006;**148**:1020–6.

91. Royal College of Surgeons. *Taskforce Report on the Impact of the European Working Time Directive*., 2014.

92. Bell RH, Biester TW, Tabuenca A *et al*. Operative Experience of Residents in US General Surgery Programs. *Ann Surg* 2009;**249**:719–24.

93. Department of Health, Chief Medical Officer, Donaldson SL. Unfinished Business: proposals for the reform of the Senior House Officer grade. 2002:0.

94. Norman G. Medical education: past, present and future. *Perspect Med Educ* 2012;**1**:6–14.

95. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. *Med Educ* 2012;**46**:636–47.

96. National Audit Office. *Healthcare across the UK: A Comparison of the NHS in England, Scotland, Wales and Northern Ireland*., 2012.

97. Van Hove PD, Tuijthof GJM, Verdaasdonk EGG *et al*. Objective assessment of technical surgical skills. *Br J Surg* 2010;**97**:972–87.

98. Grober ED, Roberts M, Shin EJ *et al*. Intraoperative assessment of technical skills on live patients using economy of hand motion: establishing learning curves of surgical competence. *Am J Surg* 2010;**199**:81–5.

99. Ackerman PL. A theory of adult intellectual development: Process,

personality, interests, and knowledge. *Intelligence* 1996;**22**:227–57.

100. Wanzel KR, Hamstra SJ, Anastakis DJ *et al*. Effect of visual-spatial ability on learning of spatially-complex surgical skills. *Lancet* 2002;**359**:230–1.

101. Mergel B. Instructional design and learning theory. *Retrieved Oct* 1998:1–29.

102. Anderson JR. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press, 1983.

103. Alaker M, Wynn GR, Arulampalam T. Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *Int J Surg* 2016;**29**:85–94.

104. Clarke P, Carty H, Tsim P *et al*. Do previous dexterity skills influence performance in single incision laparoscopic surgery (SILS)? Comparing joint force harrier pilots to medical students. *Int J Surg* 2011;**9**:499.

105. Johal K, Tsim P, Redfern A *et al*. Single-Incision Laparoscopic Surgery Versus Conventional Techniques. *Bull R Coll Surg Engl* 2012;**94**:348–50.

106. Levy PS, Lemeshow S. *Sampling of Populations: Methods and Applications: Fourth Edition*., 2011.

107. Reynolds W. The first laparoscopic cholecystectomy. *JSLS* **5**:89–94.

108. Data from Health and Social Care Information Centre (HSCIC), Hospital Episode Statistics 2013/14. 2014.

109. Sinha S, Hofman D, Stoker DL *et al*. Epidemiological study of provision of cholecystectomy in England from 2000 to 2009: retrospective analysis of Hospital Episode Statistics. *Surg Endosc* 2013;**27**:162–75.

110. Gallagher AG, Al-Akash M, Seymour NE *et al*. An ergonomic analysis of the effects of camera rotation on laparoscopic performance. *Surg Endosc Other Interv Tech* 2009;**23**:2684–91.

111. Aggarwal R, Grantcharov T, Moorthy K *et al*. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg* 2007;**245**:992–9.

112. Hance J, Aggarwal R, Moorthy K *et al*. Assessment of psychomotor skills acquisition during laparoscopic cholecystectomy courses. *Am J Surg* 2005;**190**:507–11.

113. Schijven M, Jakimowicz J. Construct validity: Experts and novices performing on the Xitact LS500 laparoscopy simulator. *Surg Endosc Other Interv Tech* 2003;**17**:803–10.

114. Tsuda S, Scott D, Doyle J *et al*. Surgical Skills Training and Simulation. *Curr Probl Surg* 2009;**46**:271–370.

115. ISCP (Intercollegiate Surgical Curriculum Programme) Procedural Based Assessment (PBA). 2010.

116. Newman I, Newman C. *Conceptual Statistics for Beginners*. Lanham, MD: University Press of America, 1994.

117. Alston C. Stanines: Definition & Explanation. 2014.

118. CAE LAPVR Surgical Simulator for Laparoscopic Training. 2014.

119. Gladwell M. *Outliers. The Story of Success*. Little, Brown and Company, 2005.

120. General Medical Council. *The State of Medical Education and Practice in the UK 2014.*, 2014.

121. Fitzgerald JEF, Tang S-W, Ravindra P *et al*. Gender-related perceptions of careers in surgery among new medical graduates: results of a cross-sectional study. *Am J Surg* 2013;**206**:112–9.

122. Bruce AN, Battista A, Plankey MW *et al*. Perceptions of gender-based discrimination during surgical training and practice. *Med Educ Online* 2015;**20**:1–10.

123. Rudnicki J, Zyśko D, Kozłowski D *et al*. The Choice of Surgical Specialization by Medical Students and Their Syncopal History. *PLoS One* 2013;**8**:4–11.

124. Lawal TA, Afolabi AO. Factors influencing the choice of surgery as a career by pre-registration interns. *Afr Health Sci* 2013;**13**:814–9.

125. Royal College of Surgeons of England. Surgical Careers. Women in Surgery. Statistics. 2013.

126. Davis EC, Risucci DA, Blair PG *et al*. Women in Surgery Residency

Programs: Evolving Trends from a National Perspective. *J Am Coll Surg* 2011;**212**:320–6.

127. RAF Cranwell. Aircrew Selection Information. Pilot. 2014.

128. Beermann J, Tetzlaff R, Bruckner T *et al*. teaching with cognition: Three-dimensional visualisation improves understanding of surgical liver anatomy. *Med Educ* 2010;**44**:936–40.

129. Goris J, Jalink MB, ten Cate Hoedemaker HO. Training basic laparoscopic skills using a custom-made video game. *Perspect Med Educ* 2014;**3**:314–8.

130. Ou Y, McGlone ER, Camm CF *et al*. Does playing video games improve laparoscopic skills? *Int J Surg* 2013;**11**:365–9.

131. Jalink MB, Goris J, Heineman E *et al*. The effects of video games on laparoscopic simulator skills. *Am J Surg* 2014;**208**:151–6.

132. Nock AJ. *The Theory of Education in the United States*. Ludwig von Mises Institute, 1931.

133. Ericson K, Krampe R, C T-R. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993;**100**:363–406.

134. Mathews G, Jones D, Chamberlain A. Predictors of individual differences in mail-coding skills and their variation with ability level. *J Appl Psychol* 1992;**77**:406–18.

135. Ackerman PL, Cianciolo AT. Cognitive, Perceptual-Speed, and Psychomotor Dterminants of Individual Differences During Skill Acquisition. *J Exp Psychol Appl* 2000;**6**:259–90.

136. Ellis H. *Famous Operations*. Marwal Publishing Co., 1984.

137. Alvand A, Auplish S, Gill H *et al*. Innate Arthroscopic Skills in Medical Students and Variation in Learning Curves. *J Bone Jt Surg* 2011;**93**, DOI: 10.2106/JBJS.K.00199.

138. Buckley CE, Kavanagh DO, Nugent E *et al*. The impact of aptitude on the learning curve for laparoscopic suturing. *Am J Surg* 2014;**207**:263–70.

139. CAE ProMis Laparoscopic Simulator.

140. Buckley CE, Nugent E, Neary PC *et al*. Do plastic surgical trainees naturally self-select based on fundamental ability? *J Plast Reconstr Aesthetic Surg* 2014;**67**:1303–4.