Education and Training in Ultrasound –

Guided Regional Anaesthesia

(UGRA)

Dr. Atif Shafqat

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

JUNE 2015

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

DEDICATION

I dedicate this thesis to my wonderful wife "Shumaila", my children "Simra" and "Shahir" and above all my "Parents" without their continuous understanding, support and encouragement, inspiration and most important; "the belief in me", my achievement would not be possible.

SUMMARY OF THE THESIS

The PhD thesis contains five chapters, i.e.: the first chapter, presenting the *"Background",* information relevant to the topic in the form of literature review, the second chapter, comprising of the first study project *"Visuospatial Ability As A Predictor Of Novice Performance In Ultrasound – Guided Regional Anaesthesia (UGRA)"*. The third chapter related to the second study project *"Construct Validity And Reliability Of An Objective Structured Assessment Tool For Performance Of Ultrasound – Guided Regional Anesthesia (UGRA)"*. This is followed by the fourth chapter encompassing the third and final study project *"The Impact of Video games on Ultrasound – Guided Regional Anaesthesia (UGRA) skills"* and finally the last chapter revealing the *"Conclusions"*.

The **first chapter** of the thesis the *"Background"* includes an introduction to the topic. It reviews the achievement of proficiency in procedural skills, common principles regarding the educational theory behind assessment. It also examines the literature about the diverse methods that can be used for the assessment of procedural skills in anaesthesia and make important implications from the literature in other fields of medicine.

In the **second chapter**, entitled *"Visuospatial Ability As A Predictor Of Novice Performance In Ultrasound – Guided Regional Anaesthesia (UGRA)"*, the experimental and original study will be presented, which intends to identify if visuospatial ability could predict technical performance of an ultrasound – guided

needle task by novice operators, and also to describe how emotional state, intelligence and fear of failure could have impact on this.

In the **third chapter**, termed *"Construct Validity And Reliability Of An Objective Structured Assessment Tool For Performance Of Ultrasound – Guided Regional Anesthesia (UGRA)"* another experimental and original study project will be validated, focusing on the construct validity and reliability of a new assessment tool for the performance of ultrasound – guided regional anaesthesia, by examining whether it can adequately differentiate between performance levels in anaesthetists across the spectrum of expertise.

In the **fourth chapter**, designated *"The Impact of Video games on Ultrasound – Guided Regional Anaesthesia (UGRA) Skills"* additional experimental, original part will be presented, focusing on whether playing video games will predict psychomotor performance of an ultrasound – guided needling task or could impact on tests of visuospatial ability and also examine the correlation between these innate abilities.

At the end of each chapter are drawn a series of conclusions, concerning the investigations performed and the results obtained. The *"Conclusions"* are summarized in **chapter five**, where the possible future investigations are also indicated.

PUBLICATIONS

Some parts of this thesis have contributed to a publication and an additional two other publications are also under consideration.

PEER-REVIEWED PUBLICATION

Shafqat, A., Ferguson, E., Thanawala, V., Bedforth, N.M., Hardman, J.G., McCahon, R.A., *Accepted for publication in Anesthesiology in June 2015*. Visuospatial ability as a predictor of novice performance in ultrasound – guided regional anesthesia (UGRA). **(Chapter 2)**

MANUSCRIPTS IN PREPARATION FOR SUBMISSION FOR PEER-REVIEW

Shafqat, A., Rafi, K.M., Thanawala, V., McCahon, R.A., Bedforth, N.M., Hardman, J.G. *Unpublished manuscript*. Construct validity and reliability of an objective structured assessment tool for performance of ultrasound – guided regional anesthesia (UGRA). **(Chapter 3)**

Shafqat, A., McCahon, R.A., Bedforth, N.M., Hardman, J.G. *Unpublished manuscript*. The impact of video games on ultrasound – guided regional anaesthesia (UGRA) skills. **(Chapter 4)**

AUTHOR CONTRIBUTIONS TO MANUSCRIPT ACCEPTED FOR PUBLICATION

Chapter 2

***Shafqat, A.**, Ferguson, E., Thanawala, V., Bedforth, N.M., Hardman, J.G., McCahon, R.A., Accepted for publication in Anesthesiology in June 2015. Visuospatial ability as a predictor of novice performance in ultrasound – guided regional anesthesia (UGRA).*

This study was truly a collaborative effort. As the primary author, I was responsible for formulating the research questions, refining the study design, conducting the study and all statistical analyses, drafting the manuscript, responding to reviewers' reports and coordinating submission and publication of the original research paper.

All co-authors were involved in planning the study; choice of measures including the statistical methods used and commented on the manuscript for submission, response to reviewers' reports and final version of the original research paper. Anonymous reviewers provided constructive comments on the submitted manuscript, which were incorporated into the manuscript.

AUTHOR CONTRIBUTIONS TO UNPUBLISHED MANUSCRIPTS

Chapter 3

Shafqat, A., McCahon, R.A., Bedforth, N.M., Hardman, J.G. Unpublished manuscript. Construct validity and reliability of an objective structured assessment tool for performance of ultrasound – guided regional anesthesia (UGRA).

This study was a team effort. As the primary author, I was responsible for formulating the research questions, refining the study design, conducting the study and all statistical analyses, drafting the manuscript. All co-authors were involved in planning the study; choice of statistical methods used and commented on the manuscript for submission. It is now in the latter stages of preparation for submission to a peer-reviewed journal.

Chapter 4

Shafqat, A., McCahon, R.A., Bedforth, N.M., Hardman, J.G. Unpublished manuscript. The impact of video games on ultrasound – guided regional anaesthesia (UGRA) skills.

The impetus for this paper arose after attending the "*N-trans sandpit meeting*", which is a translational research meeting, held twice a year for postgraduate research students in the faculty of Medicine and Health Sciences, University of Nottingham. I got inspired from a surgical presentation (Survey), which highlighted the relationship of video games with laparoscopic surgery.

This study was conceived and implemented by myself. As a primary author, I then formulated the research questions and was responsible for the overall design and conduct of the study and statistical analyses, drafting the manuscript. All co-authors reviewed and commented on the design of the study, the manuscript for submission. It is now in the latter stages of preparation for submission to a peer-reviewed journal.

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER – 1

Background

Ultrasound-Guided Regional Anaesthesia

Ultrasound is a new technology used noninvasively in regional anaesthesia to locate and block peripheral nerves. It offers a number of clinical benefits. It permits clear visualization and localization of nerves.[1-6] As a result, this technology assists in identifying the best possible site of insertion for the block needle.[7] Ultrasound also provides real-time guidance for the block needle during its advancement toward the target nerve, which assists in decreasing random needle movements. At the time of injection the spread of the local anaesthetic can also be detected which helps to make sure that an adequate amount of drug has been deposited completely around the target nerve.[2,3,5,8,9] Ultrasound-guided regional anaesthesia also reduces complications including accidental intravascular or intraneural injection.[1-3,5,9-11] According to Griffin J et al. the usage of ultrasound in regional anaesthesia has become widespread because it enabled the operator to avoid using the classically described anatomical landmark techniques.[12]

Status Of Ultrasound-Guided Regional Anaesthesia

Ultrasound-guided regional anaesthesia is a significant and rapidly developing practice in regional anaesthesia.[7] The practice of ultrasound-guided regional anaesthesia has increased extensively due to its improved efficacy and safety.[13-18] There is a substantial amount of evidence available, which reveals that ultrasound-guided peripheral nerve block when compared with peripheral nerve block via nerve stimulation increases block efficacy, prolongs duration of the block, shortens block performance time and above all improves patient safety.[11,13-15,19] Ramlogan R et al.[20] surveyed the members of the American Society of Regional Anaesthesia (ASRA). The

survey illustrated that 67% of 583 members used ultrasound-guided regional anaesthesia. There is a strong consensus among different experts that ultrasound practice will soon become the gold standard for regional anaesthesia.[1,19,21]

Medical education in the United States has changed from an apprenticeship model of learning to competency-based training. During this time the practice of regional anaesthesia has also advanced – from a period of 'clicks' and 'pops' and paresthesias to stimulating catheters and ultrasound guidance. The training curriculum in regional anaesthesia is still based on the "see one, do one teach one" philosophy.[22]

There are two elements of learning ultrasound-guided regional anaesthesia. First, the interpretation of sono-anatomy that involves identification of anatomic structures as seen on the ultrasound images, and second the needling technique, which require learning to control the ultrasound transducer and needle to guide the needle to the target under direct vision. This needling technique requires knowledge of sono-anatomy and hand-eye coordination. Ultrasound-guided regional anaesthesia is analogous to complex surgical tasks and therefore may be successfully trained if broken down into component parts and consecutively learned before integrating them into complete performance of the procedure.[23] American Society of Regional Anaesthesia (ASRA) and European Society of Regional Anaesthesia (ESRA) joint committee recommendations for education and training in ultrasound-guided regional anaesthesia put forward a number of useful resources; a list of 10 essential tasks in performing ultrasound-guided regional anaesthesia, recommended training routes, a description of the core competencies, and a recommended curriculum for training in

ultrasound.[24] Thus, these guidelines provide a method by which physicians can become educated in ultrasound-guided regional anaesthesia.

Present Training Model

The health industry lags behind the other industries performing dangerous work. They have not been able to guarantee that their human resources consist of employees who are consistently and provably competent practitioners.[25,26] The Institute of Medicine published a report in 1999.[27] According to that report medical errors have been described as the eighth most important cause of death in the US. As a result of this report the standard of healthcare and patient safety have become the centre of attention of the medical profession. Training and inexperienced physicians due to their deficiency of knowledge and technical skills can compromise patients' safety and care.[28] They are generally exposed to limited resources and commonly encounter inadequate supervision by the more experienced physicians. In the acute settings that include the intensive care unit, the physicians may also have inadequate experience of non-technical skills like teamwork, communication, leadership and management.[29]

Fatigue is among one of the other factors that can play a crucial role in the performance of the physician and may cause medical errors.[30-33] Work hour limitations have been put into practice by Accreditation Council for Graduate Medical Education (ACGME), in order to counteract the fatigue problem and for the improvement of the patient safety. The study by Landrigan CP et al.[34] demonstrated a decrease trend of medical errors in the intensive care unit after the work hour limitations. But the trend of medical errors including the procedure related errors made by 'rested' house officers continued to remain high. In another study, Barger LK

et al.[35] showed increased trend of significant medical errors, adverse events and attention failures by those physicians who worked extended-duration work shifts. Thus, the above studies provide evidence that in addition to fatigue, deficiency of knowledge, skills and competency is responsible for causing medical errors by the physicians.

The existing structure of medical training is based on the apprenticeship model of 'learning by doing'. This system is also commonly known as the 'see one, do one, teach one' model.[22] According to this system the physician gains experience during training which assist them in learning and at the same time encourage them to become self-reliant and independent. This long-established model is deficient for attaining the essential components of competency in medicine. This includes constant supervision, assessment of performance and feedback by an organized and structured technique. According to the existing model the physician becomes competent after performing a precise number of procedures or by completing a set number of postgraduate years of training.

There is no official assessment of any competencies of procedural skills in most of the countries around the world, including the US. The assessments usually occur either after or towards the end of the physician training years. Therefore physicians continue practicing independently without being assessed for any official competencies. According to Reznick RK et al.[36] physicians should undergo competency-based assessments to acquire promotion and advancement in the training programme. This method of promotion and advancement should be preferred over a specific time period spent in the training and be included in the current certification processes.

Many different methods have been suggested and commenced to develop the present training programme. Some of the methods have a characteristic framework for safety culture.[37] According to Sexton JB et al.[38] and Thomas EJ et al.[39]; this safety framework can be evaluated and appraised with extensively used validated tools and is coupled to the clinical and operational outcomes. Volpp KG et al.[40] have analyzed and recommended a number of simple and economical ways to decrease medical errors in hospitals. They also proposed to replace this current training model with a rational system for training physicians to perform procedures.

Drawbacks Of The Present Training Model

Physicians in hospitals commonly care for critically ill patients in acute specialties like the emergency department, operating theatres and the intensive care unit. These patients in all of the acute settings mentioned above are at risk and therefore any mistakes made by any of the physicians in their care pathway can lead to a disastrous and irreversible outcome.[41-46] In addition to the deficiencies in education and training, there are other causes of medical errors. These involve ineffective communication and teamwork, inadequate supervision and failure to implement protocols or total absence of recognized protocols.[47-49] A study by Baldwin PJ et al.[50] demonstrated that the residents made a considerable amount of medical errors. These were more often due to inadequate knowledge and lack of experience, which suggests that physicians need closer supervision in the early years of their training. In two additional studies, between 28% and 42% of the residents feel that they are not sufficiently trained enough to carry out a medical procedure safely and without any assistance for the first time[51,52]. Similarly, multiple simulation studies[53-55] have revealed that inadequate knowledge, lack of experience and inadequate supervision are the main causes of

medical errors. Cox CE et al.[56] revealed that 29% of the internal medicine residents feel that they are inadequately trained to look after mechanically ventilated patients. Almost half of them towards the end of their residency still believed that they are incompetent to manage those patients. Wu AW et al.[30] discovered in his study that 45% of the internal medicine house officers have reported that they make at least one error, in which 31% of the cases end up as fatal.

Typically, one type of error that is quite common in physicians is a procedural error. The procedural errors can be both diagnostic and therapeutic. There is relatively insignificant epidemiological data regarding procedural errors. There are two main reasons for this little data. Firstly, the errors are under-reported[27] and secondly, most of the data that is available is obtained self reportedly, which considerably underestimate the significance of errors.[57] The results of Harvard Medical Practice Study I & II illustrate that the second most recurrent cause of adverse events are procedural and therapeutic errors.[46,58] Similarly, in another recent study by Jagsi R et al.[28] it was proven that 31% of the most frequent medical errors that were reported by the physicians, were accountable to procedures out of which 25% of those were fatal or life threatening. According to the Critical Care Safety Study Group, 80% of performance failures which consists of adverse events and medical errors were due to skill-based slips and lapses and also knowledge- based mistakes.[59] Patient safety reporting software has also highlighted the problems of procedural errors. The intensive care unit (ICU) safety reporting system (ICUSRS) study[60-62] was a prospective study of incidents reported by nurses and physicians. Forty nine percent (49%) of the incidents recorded illustrated deficiencies in training and education.[60] To sum up, the above evidence implies that physicians are inadequately trained to carry

out procedures. These physicians with not enough experience and little supervision can be unsafe for patients.

A New Training Model

The development of education for health care professionals is crucial for the improvement of the quality and standards of health care. The aim is to develop an outcome-based education approach that can enhance and groom clinicians to encounter both the requirements of the patients and the changing health system.[57] There is a need for change in the training of physicians from giving emphasis on the training process rather than the examination results. This shift from input-based education to results-oriented education has been highlighted in recent years.[63]

There are multiple organizations that have defined their scopes of practice for the use of this new outcome-based approach. The Accreditation Council for Graduate Medical Education (ACGME) in the USA has recognized the outcomes for graduate medical education.[64] According to Schwarz and Wojtczak[65], the Institute for International Medical Education (IIME) has recognized and implemented the minimum international prerequisites in medical education. The General Medical Council (GMC) in the UK has also declared its stance to concentrate on outcome-based approach.[66] The Scottish Deans' Medical Curriculum Working group in the year 2000 has forwarded their recommendation to focus and implement on the outcome-based model with 12 learning outcomes, to be set up in all of their Scottish medical schools.[67] The UK Department of Health[68], put forwarded the proposals on the transformation of the medical training in the UK. As a result, the Medical Royal Colleges in the UK started to classify the curriculum for postgraduate training on outcome-based model. The

Joint Committee on Higher Medical Training[69], then revealed a standard basic and nonspecific higher medical training curriculum, which was based on the General Medical Council's good medical practice outcomes.[66] In the same way, The Joint Committee on Higher Surgical Training classified and categorized the essentials of surgical competence.[70] The World Health Organization (WHO) in 1978 published an essential report on the significance of competency-based model.[71] According to the report competency-based curriculum is vital for the education of the health care professionals. It also predicted that this model would match with the service needs of the National health. It further highlighted the importance of the assessment of the defined competencies as the major determinant of any training package success. According to the institute of medicine report in 2003, all health professionals regardless of their discipline must possess core a set of competencies and therefore should be educated and trained to deliver patient-centered care as members of an interdisciplinary team. The report has emphasized about evidence-based practice with quality improving approaches and the use of informatics.[57]

One of the required characteristics of competence-based approach to education is to be able to assess the physicians' accomplishments of the basic competencies or outcomes. The learning outcomes and the methods of assessment by which the physicians' accomplishments are appraised should be listed as clearly as possible. The terms 'competence' and 'performance' are interchangeable. Rethans et al.[72] have described the difference between competence-based and performance-based assessment. In competence-based assessment we appraise the physician in controlled representation of professional practice whereas in performance-based assessment we evaluate the physician in their professional practice. Competence is dependent on

circumstances and situation. This is because it correlates the capability of the health care professional, the task that needs to be completed, the working environment, and the clinical contexts in which the task is performed.[73] Therefore competence can only be assessed in the work place. The professionals should be familiar with what is required in order to provide safe care. They should know how to use their knowledge, which they have accumulated.

The safety of the patients today is more and more recognized as an essential measurement of quality care. As a result it has been included gradually in the education curriculum of healthcare professionals.[74-77] There is a significant amount of evidence available, which consists of studies on the assessment of the competency of health care professionals but none of the previous studies have focused on the aspect of patient safety.[78,79] Hence professional bodies and the World Health Organisation (WHO) have developed and sanctioned a patient safety competencies framework for healthcare professionals. The main purpose of this framework is to develop and improve local patient safety training programs.[16,57,80,81] Thus, preferably, there would be a toolbox available that contains all of the essential elements of the defined patient safety competencies throughout the curriculum. Each of these assessments will be made at that level of Miller that is appropriate to the level of training achieved by the professional.[82]

The current medical training is inadequate particularly for technical and procedural skills. Due to this inadequacy of current training, patients' wellbeing is being compromised. Rodriguez-Paz JM et al.[83] proposed a new and innovative training model that should combine knowledge and skills-based learning. The teaching should

have emphasis on patients' safety in both simulated and actual patients' settings. In this model of training the physicians should be evaluated for their knowledge, skillfulness and be tested for their approach towards patient safety. This training model should be based on the four principles of Miller GE's framework for the assessment of clinical competence.[82] Richardson V[84], and Schon DA[85] have recommended that this model should transform the process of 'action-reaction' to the process of 'action-reflection'. They believe that this new approach of 'action-reflection' is more realistic and it resembles very closely to the manner by which humans learn. Instead of including numerous competencies related to the practice of the physicians, Rodriguez-Paz JM et al.[83] recommended inclusion of deficiencies and errors in the development of the curriculum. Learning, repeated practice, performance and evaluation form the key components of this training model. Within this model the physicians are taught that mistakes are associated with medical training. The physicians practice new skills in a safe and simulated medical setting. Thus they are able to act and reflect on their skills before they actually carry out those skills in actual patients.

Competence And Assessment

During the mid-1960s, competency-based training began in the USA. At that time the quality of teaching in schools was not up to standard and was criticized. Many considered that the curriculum for teacher training had little correlation to the job requirements of teaching.[86] Then with the assistance of federal government funding, several colleges began to implement on this new educational model and started teaching and training student teachers according to a new competency-based curriculum. The curriculum focused on classroom skills. The assessment was then based on practice and concentrated more towards observation of skills in action.[87]

In the UK, during the early-1980s there was an enormous increase in youth unemployment. Due to that, many employers at that time were of the opinion that school leavers were deficient with the type of skills needed for work. The vocational education in the UK was compared critically with other developed countries. As a result, the government was convinced about the education reforms and started thinking on the introduction of competency-based training. The main objective of the vocational courses is to teach skills applicable to the relevant employment. Many experts have extended this main objective to include learning in terms of changes in behavior. These changes are known as learning outcomes. These changes are defined in a new type of syllabus, which concentrates on the skills, knowledge and understanding the learner will attain at the end of the course rather than listing what is to be learned. The new vocational curriculum lists learning aims in great detail. The competence model of the curriculum is related to a new view of evaluation. Conventionally, a test is said to be successful if it consistently measures a candidate's performance in terms of scores in the provision of a pass mark. According to this new

view a test is only successful when the pass mark guarantees the good accomplishment of the candidate by application of the knowledge attained into the practical work. This type of assessment cannot be implemented in examination centres and hence observations of the candidate in action have a vital role in the assessment process.

In medicine most examinations intend to look at the deep understanding of fundamental concepts of the professional. They do not aim to test and analyze the capability to do a job. Recent trends have shown that there is a shift within general education towards learning and assessment that are more focused on the ability of the learner to do the job. These changes towards competency-based education are essential enough to cause a paradigm shift in education. In this innovative educational model, assessment is done by means of direct observation of competence. According to Jessop G[88], "Assessment of performance in the course of normal work offers the most natural form of evidence of competence and has several advantages, both technical and economic." Medical education is vocational. Medical curriculum has typically been academic in nature with conventional exams towards the end of the year. These exams are used as a mode of assessment and promotion. Medical graduates are being taught both theory and clinical practice, which have been incorporated into to the curriculum to make them clinically competent professionals. New methods of teaching medical graduates have preferred learning outcomes than success in academic examinations. Problem based learning is a form of a performance-based approach to education, in which the curriculum improves the skills of medical practice. These skills include the use of decision-making and clinical reasoning. Learning and assessments are closely connected to each other and

education is holistic. The use of both competency-based teaching and assessment is now being encouraged in medical education.[89]

An assessment is something we carry out to evaluate the outcome of a learning experience. It can acquire various different forms such as an examination, a multiple-choice questionnaire (MCQ), an oral examination, an objectively structured clinical examination (OSCE) or a workplace observation. Assessment is a necessary element of competency-based education.[90-92] The main usage of assessment should be to provide feedback and for summative examination. Every assessment should be in accordance with a simple set of rules. They should have a purpose and an outcome. In addition, it is very important that they must be valid and reliable. A valid assessment is one that measures what it is supposed to measure. A reliable test is one that produces the same results if the candidate immediately retakes the test and irrespective of which examiner marks it. An assessment also expresses something about the philosophy of education and learning. Traditional academic education is associated with the view that learning trained the mind in ways that were of general benefit to future performance. Cardinal Newman (1852)[93], expressed this view of liberal education as: "*Even though it be turned to no further account, nor subserve any direct end…. To open the mind, to refine it, to give it powers over its own facilities is an object as intelligible as the pursuit of virtue.*"

Assessment Of Clinical Competency

One of the most difficult challenges confronted by medicine today is the assessment of clinical competence. Miller GE[82] has provided a framework in the shape of a pyramid for the assessment of clinical competence [figure 1]. This pyramid framework consists of 4 levels. The first level is termed as 'knows', which is the 'knowledge' base that is considered necessary by a physician in order to perform their professional job effectively. Most of the examination systems consider measurement of this knowledge base through objective test methods, to be sufficient. This first level is vital but an incomplete tool for assessment of competence. The second level is known as 'knows how', which is the capability of the physician to use the knowledge that they have acquired. The physician should develop the skill of gaining information from a range of human or laboratory sources and then should be able to carry out the analysis and understanding of the data. Last but not least, is to be able to finally convert the results into a logical diagnostic or management plan. Webster calls this logistic approach of having an adequate knowledge, skills, judgment or strength for a particular task as 'competence'. The third level is recognized as 'shows how'. There has always been a concern that what will the physicians do when they will encounter real patients? Academic examinations have failed to record this. This means that in addition to 'knows' and 'knows how', the physicians should be able to 'shows how' they do it. Clinical teachers nowadays concentrate more on this assessment of 'performance' goal. The fourth and final level is documented as 'does'. By this we mean that what a physician actually 'does' when performing his professional duties independently in a clinical practice, after undergoing assessments of all the three levels mentioned above. Thus, this is the 'action' element of the professional behavior

and on the whole is the most complex and complicated to be assessed consistently and correctly.
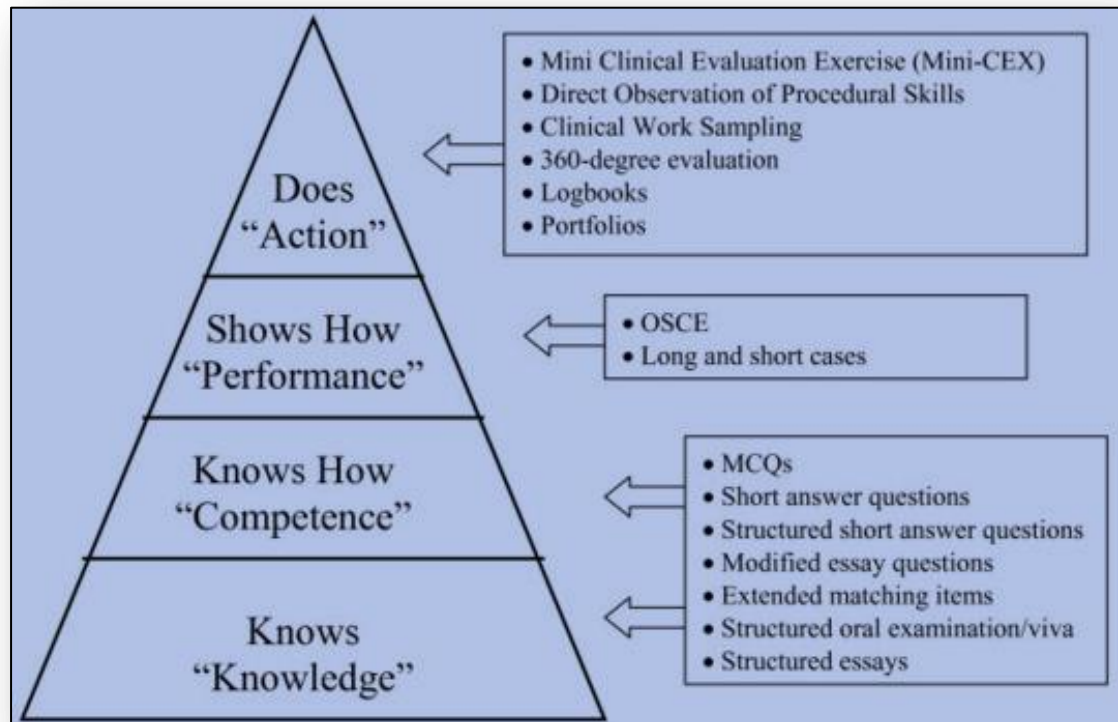
Figure 1 Framework for Clinical Assessment[82]

The first and second levels 'knows' and 'knows how' respectively, are appraised and measured by the conventional assessment tools of oral and written tests. The third level 'shows how' are assessed by practical examinations, observation of long or short cases, or objective structured clinical examination (OSCE). The fourth level 'does' can be evaluated by observing the physician at its workplace.[94]

Competency And Anaesthesia

In anaesthetic education both competency teaching and assessment are being implemented. The curriculum "Objectives of Training in Anaesthesia" by the Australian and New Zealand College of Anaesthetists is prepared on the basis of competency model.[95] On the other hand, the anaesthetic syllabus by the Royal College of the Anaesthetists[47] is not entirely based on the competency model. This competence-based learning is coupled with workplace assessments. The national examinations aim to make sure that the new medical professionals entering into the specialist practice have attained the competent status. The exams concentrate mainly on the knowledge component of the specialty, which will be tested by fair, reliable and feasible assessments when applied to large groups of candidates.[96] Practical skills, performance and personal behavior are vital for an effective anaesthetist but they are not open to formal examination testing.

Although the research data is limited to support the connection between academic accomplishment and success in the workplace, the government and the public associate examination passes with distinction and quality. Assessment at workplace consists of taking a broad view from an observation, which has been recorded during working hours. The sample of professional behavior observed must be a symbol of all related behaviors.[97,98] According to the Standing Committee on Postgraduate Medical and Dental Education (SCOPME)[99], there is a need for service-based assessment. This assessment must be impartial and supposed to be a formal process with a right of appeal. It should form one of the building blocks of the learning development. Limen referencing can provide the foundation for the standard of a workplace-based assessment.[99] According to the limen referencing the standard of the performance is

correlated to the assessor's expert understanding of the minimum standards required for the safety of the performance. Paget NS et al.[100] have considered limen-referencing approach very useful and reliable. Jolly B et al.[101] in the good assessment guide have considered this method as one of the best for assessing the performance of the complete task.

Technical Skills In Anaesthesia

The ability to perform practical procedures efficiently and safely is a significant aspect of the practice of Anaesthesia. In the specialty of anaesthesia procedural skills are given less importance and hence, are appraised inadequately when compared with other areas of learning like knowledge and judgment-based skills. Surgical outcome is associated with good quality of procedural skills.[102] In surgery there has always been an increase emphasis on this area of learning. In fact the field of surgery has lead the way in the research and the development of the assessment of procedural skills.[36]

Technical skills Accomplishment

The process of procedural skills achievement consists of three stages: cognition, integration and automation.[103] *Cognition* means development of an understanding of the task. It also includes perceptual knowledge. It is supported by a clear explanation and demonstration of the task. The *Integration* stage consists of the incorporation of the knowledge gained in cognition stage into the learning of the motor skills for that specific task. As a consequence the task becomes *automated* that is involuntary and spontaneous. In order to become an expert or be competent in a procedural skill, an individual needs to undergo and practice that task or skill multiple times, depending

upon that specific task. The standard of teaching and the talent of an individual also play an important part in achieving competence. In the anaesthetic curriculum the skills that are taught initially to the trainees are typically simple and uncomplicated. After achieving competencies in those simple and uncomplicated cases, the trainees are then introduced to a broader and advance range of normal and pathological cases, which are required in order to achieve competency and expertise. According to Hamdorf JM et al.[104] the retention of these motor skills mostly relies upon the level, to which the skill was learned, experienced and perfected.

Types of Assessment

Assessment differs from evaluation by the practice of reliable, quantitative tools with a assessable level of objectivity.[105] Assessment can be performed by the following approaches below:

Summative Assessment

Summative assessment is used to make outcome decisions.[106,107] It is completed towards the end of a training period. According to Gardner J[108], this type of assessment generally gives a grade or a pass/fail. It has generally been described as 'assessment of learning'. Summative assessments can be used for many different purposes, such as;

1) It can form a part of the progress within a competency-based training system.

2) It may possibly be used for certification or revalidation of medical licensing.

3) It can be required for promotion of an individual from junior to senior level, so that they should be allowed to have considerable level of responsibilities.

In the specialty of anaesthesia self-reported procedure lists or logbooks, retrospective subjective feedback from an educational supervisor or both has been used in the summative assessment of procedural skills.[109]

Formative Assessment

In Formative assessment information about an individual is collected and then used to provide feedback and encourage learning.[107] It is used to facilitate learning and have been described as 'assessment for learning'. This assessment is generally carried out during supervised clinical work. The feedback should be provided well in time so that it can of benefit in the individual's development. Turnbull J et al.[110] have recommended that feedback should be given as soon as possible, sometimes concurrently during the performance of the procedure. On the contrary, Schmidt RA et al.[111] and Xeroulis GJ et al.[112] have revealed the importance of feedback after the completion of the performance of the procedure. They believe that this type of feedback also facilitate in long-term retention of skills.

Quality Of Assessment

In order to be effective, assessment tools must accomplish adequate levels of performance for six characteristics including reliability and validity.[110] Feasibility means that whether an assessment tool can be used in any specific program.

Reliability

Reliability is described as the reproducibility of a test result. Two different raters independently appraise performance and accomplish similar conclusions. In education they are also known as inter-rater agreement or test-retest agreement. These agreements are frequently described as 'external reliability'. Reliability can be further illustrated by different statistical analyses below:

1) Pearson's product-moment correlation coefficient (r):

This has been used to describe inter-rater agreement. An 'r' > 0.75 means excellent agreement.[106] However, this method has been criticized by Hunt RJ[113] as not accounting for bias. For example, if one examiner consistently uses the low marks in a scale and another examiner consistently uses the high marks, they could still have a high correlation coefficient if they rank subjects similarly. This bias is particularly significant if the assessment tool has a set pass mark.

The Intraclass correlation coefficient is also used to depict inter-rater agreement. It explains the agreement that would be seen by chance. It is defined as the ratio of variance between subjects due to error variance. According to Vassiliou MC et al.[114], an ICC of 0.8 means that 80% of the variance among scores can be attributed to true variance among subjects. Muller R et al.[115] have described another type of ICC called Cohen's 'κ' coefficient. This is used when there are two raters. According to Landis JR et al. a 'κ' > 0.80 indicates a near-perfect agreement; 0.61 − 0.80, substantial agreement; 0.41 − 0.60, moderate agreement; 0.21 − 0.40, fair agreement; 0.00 − 0.20, slight agreement; and < 0.00, poor agreement.[116] This classification is not accepted universally and in a practical world the degree of acceptable agreement depends on the circumstances.

Internal consistency is not whether a subject performs similarly in each part of a test, but whether different subjects tend to do well or badly on the same parts of the test. Internal consistency is a commonly used measure of reliability and it is sometimes also referred to as 'Internal reliability'. Cronbach's 'α' gives a value ranging from 0.0 − 1.0. By rule 0.0 − 0.5 can be rated as imprecise, 0.5 − 0.8 as moderately reliable and 0.8 − 1.0 can be used with confidence for certification purposes.

Validity

Validity refers to whether the test is actually measuring what it is designed to measure. It has following types;

1) Face Validity:

Face validity is described as a general impression of whether the evaluation seems to be appropriate. For instance, evaluating performance of epidural anaesthesia on a banana model has little face validity as compared to a direct observation of the placement of an epidural in a patient.

2) Content Validity:

Content validity denotes to whether an assessment tests the content either of what was being taught or appropriate content as defined by a group of experts.

3) Concurrent Validity:

Concurrent validity determines validity based on agreement with another established valid measure.

4) Construct Validity:

Construct validity is used when there is no recognized gold standard for comparison. An education evaluation is considered valid if it can differentiate between groups with different levels of experience. It becomes progressively more valid if the groups it can distinguish between are more similar in experience.

5) Predictive Validity:

Predictive validity is the ability of a test to predict something that happens after the test such as a clinical outcome or the future test results of a subject.

Methods For Evaluation Of Technical Skills

Checklists

Conventionally, a consultant carries out the assessment of procedural skills by direct observation without criteria. Direct observation has high feasibility in the specialty of anaesthesia due to the nature of the specialty involving high level of supervision.[107] Unfortunately these assessments without a certain criteria have poor reliability and validity.

Checklists are utilized for grading performance during direct observation with criteria. Checklists split a procedural task into its component parts. A dichotomous pass/fail outcome is then allocated to each component part. Friedman Z et al.[117] have modified the checklists by removing binary pass/fail outcome and introducing outcomes of 'not performed', 'performed poorly', and 'performed well'. This has made the checklists more qualitative at the expense of losing its objectivity. Checklists have a possible drawback that is, if all phases of the task are considered equivalent regardless of any clinical significance, then the individual can achieve high score in spite of leaving out important phase. This drawback can be compensated in two ways. Firstly, certain important phases that have clinical significance can be labelled as critical phase such that in case of leaving out may result in an automatic fail. Secondly, an overall pass/fail option is added to the scoring system. The study by Friedman Z et al.[117] revealed excellent reliability of checklists in the assessment of epidural anaesthesia. Construct validity has also been established for checklists in the same study. Similarly, Naik VN et al.[106] proved that checklists when used by trained assessors have been found to have good reliability for the assessment of interscalene brachial plexus blocks. They also established the construct validity.

Global Rating Scales (GRS)

Global rating scales (GRS) make use of a Likert scale instead of a dichotomous outcome. It consists of different degrees of responses in each category. This arrangement makes them more qualitative but less objective than checklists.[114] Naik VN et al.[106] and Thompson WG et al.[118] respectively have revealed during assessment of procedural skill that global rating scale may either illustrate an overall idea of the quality of performance or there may be a Likert scale for a number of different domains within an overall performance. Global rating scale can be employed both prospectively and retrospectively but evidence revealed that when used retrospectively, reliability is compromised.[118] Like checklists global rating scales also have some drawbacks. One of the phenomenon is 'halo effect'. In this when a good or bad quality performance in one area unnecessarily influences the grade of performance in other domains. Studies have held inexperienced assessors accountable for this phenomenon.[118,119] Self-imposed scale limitation is another problem of global rating scale. The assessors limit themselves only to the high end of the scale. In fact Thompson WG et al.[118] have shown in their study that 95.6% of the scores on a nine-point scale were between 6 and 9. Again the causes for this are lack of trained assessors and/or reluctance of the assessors to fail the subject. The University of Toronto developed the first global rating scale for the assessment of procedural skills in surgery.[120,121] Construct validity of this tool has been established numerous times for the assessment of procedural skills both in surgery[120,121] and anaesthesia.[106,117] In addition, multiple anaesthetic studies[106,117,122] have also revealed good reliability for this tool in the assessment of procedural skills. At present, global rating scales are also being employed for in-training assessment of procedural skills in the UK Foundation Programme. One of the compulsory competence assessment tool is Direct Observation

of Procedural Skills (DOPS).[80] DOPS is a specific 6-point, 11 domains global rating scale that is used to evaluate performance in procedural skills.

Relationship between GRS and Checklists

Goff BA et al.[123] and Morgan PJ et al.[124] have established good reliability for both global rating scales and checklists, in their studies. In a simulated surgical study, Regehr G et al.[125] compared global rating scale and checklists and established that both tools illustrated construct validity and good reliability but global rating scale depicted better construct validity than checklists. Aggarwal R et al.[126] recommended that procedure-specific scales are much more advantageous since they provide a further degree of formative feedback to generic scales. Sites BD et al.[127] have confirmed this in their study of ultrasound-guided regional anaesthesia, where a generic scale is not suitable for a specific skill of ultrasound-guided regional anaesthesia.

Simulation

Simulation consists of the practice of manikins, human cadavers, animals, virtual reality, and standardized patients. 'Part-task' trainers (PTT) simulate a particular anatomical area or procedure as compared to full-patient simulation with a computer-enhance mannequin. PTT are usually less costly than full-patient simulators. According to Vadodaria BS et al.[128] full-patient simulators deliver better results in terms of reconstructing the whole clinical situation. By fidelity of a simulation we mean that how much closer is the simulation task to the 'real world' or actual task. There are numerous advantages of appraising procedural skills in simulated conditions rather than in actual clinical conditions. According to Ziv A et al.[129] assessments of procedural skills in a simulated setting avoid possible injury to patients. They have described it as an ethical imperative. Anaesthetic simulations like the Bill 1 airway simulator for cricothyrotomy[130] and a virtual reality flexible bronchoscopy simulator[131] have all established their construct validity.

Cumulative Sum Analysis (Cusum)

Cumulative sum analysis was originally developed in industry for quality control. It is a statistical method that focuses on the outcome instead of the process of performance of the procedural skills. It plots a graph of the subject's performance over time based on predetermined criteria for success or failure. A value for Cusum is plotted on the y-axis and the number of attempts on the x-axis. Failures result in a move up (and successes down) the y-axis as the subject progresses through increasing numbers of attempts. When the Cusum score decreases below a level based on a predetermined 'acceptable failure rate', the subject can be considered to be competent with statistical significance. The distance of the individual from a predetermined line indicates that how far they are from achieving competency as defined by Cusum.[132] The other endpoints such as a change in the curve refer to an improvement or decline in the performance.[132,133] The Cusum analysis is an effective objective tool to describe learning curve. The number of cases or attempts required in order to attain competency can be projected by the usage of Cusum analysis.[133] According to Naik VN et al.[134] the required number of cases or attempts differs extensively between different individuals. The study by de Oliveira Filho GR[135] revealed that the number of attempts to achieve competency in tracheal intubation was between 09 and 88. One of the weaknesses of the Cusum analysis is to be dependent on self-reporting that can make it incorrect. The objectivity of the Cusum depends upon the pre-determined success or failure endpoints. The evidence proves that the classification of the endpoints can be different extensively.[133,135,136]

Logbooks or Procedural Lists

Traditionally, procedural skills have been assessed by a combination of a subjective opinion from an educational supervisor and logbooks or procedural lists.[109] Due to high feasibility procedural lists have retained their popularity and are still being used as a common type of assessment of procedural skills. In order to advance and undergo development of acquisition of procedural skills, one must perform a specific number of procedures.[131,135] This specific number of procedures is highly variable between individuals. Procedural lists are used for summative assessment. Their usefulness is limited when used for summative assessment. The reason for this limitation is because there is no assurance of correct performance of the procedure by the individual. The individuals can repeatedly make mistakes and at the same time regard their own performances satisfactory. According to Bould MD et al.[137] procedural lists are excellent tool for assessing the training prospect of the posts or the programme rather than for assessing an individual.

Psychometric and aptitude testing

The role of Psychometric testing in the surgical field for predicting procedural performance has little significance. Wanzel KR et al.[138] in his study have found mixed results. He revealed statistically significant and modest correlation between the Z-plasty procedure performance and its scores in tests that assesses the ability to rotate 2D and 3D figures mentally. He did not find any correlation with less complex tests that assess the recognition of simple shapes. Similarly, according to Gallagher AG and colleagues[139] there was a modest correlation on comparison between performance in PicSOr and laparoscopic performance of both novices and expert laparoscopic surgeons. The MICRO computerized Personnel Aptitude Tester (MICROPAT) measures psychomotor ability. It has been investigated in different anaesthetic studies[140,141] and has been found of limited application. Psychometric testing in anesthesia at present is essentially a research tool and mostly unproven.

Other Methods for Evaluation of Technical Skills

The Objective Structured Assessment of Technical Skills (OSATS):

OSATS is related to the objective structured clinical examination (OSCE). After the widespread recognition of OSCE, a group in Toronto developed an analogous model for the appraisal of technical skills.[120] According to Reznick R et al.[121] it is an objective assessment of procedural skill outside the operating theatre. It consists of six stations where residents and trainees perform procedure tasks on live animal or bench models in fixed time periods. During the performance of the tasks the residents and trainees are assessed by procedures specific checklists and global rating scale. Multiple studies[120,121,123,142] have provided evidence regarding OSATS to be a reliable instrument for measurement of procedural skills. OSATS have also established its constructive[121,123] and concurrent validity. The drawback of OSATS is that they are relatively expensive to run and difficult to put in order because of its extensive setting.[143] By using this method of assessment Datta V et al. demonstrated the construct validity of the global rating scale with an inter-rater reliability of 0.81.[144]

The Imperial College Surgical Assessment Device (ICSAD):

ICSAD is a commercially available electromagnetic tracking system (Isotrak II, Polhemus, United States). This motion analysis device originally intended to investigate the hand movements of the surgeons. The construct validity of ICASD for a range of laparoscopic and open surgical tasks has already been established.[145] Nowadays it has started making its way into the specialty of anaesthesia.[146]

CHAPTER – 2

Visuospatial Ability As A Predictor Of Novice Performance

In Ultrasound – Guided Regional Anaesthesia (UGRA)

INTRODUCTION

Opportunities to develop practical expertise in contemporary United Kingdom medical practice continue to be eroded by the work time regulations, an increasing focus on service delivery, and the need to ensure patient safety. It is no longer acceptable to allow medical trainees to practice on an unsuspecting public. Thus, a need exists to accelerate expertise development.

The ability to perform practical procedures competently is essential to the safe practice of anesthesia. Ultrasound–guided regional anesthesia (UGRA) is a complex, invasive procedural skill that requires manual dexterity, hand-eye coordination and a working knowledge of sono-anatomy.[147] International regional anesthesia societies have emphasized the need for training and competency assessment in UGRA to ensure safe practice.[24,148,149] However, some trainees will learn more quickly than others.[150] Early identification of those who may require additional support is key to developing efficient expertise acquisition within the time constraints of postgraduate training.

The assessment of procedural skills is anesthesia has generally been given less importance than the assessment of knowledge and judgment-based skill.[107,151] Anesthesia is predominantly a procedural specialty but has lagged behind other domains[152-159] in assessing the technical skills of its professionals. Visuospatial ability is the capacity to mentally rotate and manipulate 2D and 3D objects. Previous work in laparoscopic and endoscopic surgery has demonstrated that visuospatial ability correlates significantly with the performance.[138,154,155,159-167] Mental rotation is a

visuospatial ability, which correlates positively with novice performance of simple laparoscopic tasks on bench-top models.[159,160,166,168] At a basic level, it is possible that laparoscopy and UGRA are similar with respect to the interaction of the operator's hands and eyes with the ultrasound probe/laparoscope, the patient and the screen. The importance of visuospatial aptitude has been emphasized in UGRA skills acquisition.[147] However, there is little evidence supporting the role of visuospatial testing in UGRA. The primary aim of this study was to determine whether visuospatial ability could predict novice performance of an ultrasound-guided needle task. Specifically, we chose to study the Mental Rotation Test (MRT), the Group Embedded Figures Test (GEFT), and the Alice Heim Group Ability Test (AH4).

Furthermore, procedures undertaken both in the real clinical context and while being assessed are likely to generate emotional responses, such as anxiety associated with the task. All these are known to have detrimental effects on the performance.[169-172] The anxiety relationship of the participants is previously known to influence how well they perform the task. This anxiety-performance relationship may be linear (higher anxiety is linked to reduced performance) or quadratic, represented by an inverted U-shaped function (that is low and high anxiety are linked to reduced performance with greatest performance at optimal levels of anxiety).[170,173] The trait, fear of failure is a sub-clinical form of state anxiety when success is being valued. Fear of failure is seen as synonymous with anxiety.[174-176] At a dispositional level, traits such a fear of failure and competence beliefs that people hold, could potentially hinder performance by enhancing the negative effects of anxiety (fear of failure increases anxiety that eventually leads to performance deficit – a simple mediation model).[177,178] In addition,

general cognitive ability is considered to be one of the best predictors of performance overall.[179]

Therefore, as well as examining the effect of cognitive skills on the ability to perform procedures like this, the effect of emotional processes also needs to be explored. Unfortunately most of the previous studies have focused exclusively on cognitive skills.[138,147,155,161,167] This study therefore, adds to the literature by also including assessments of anxiety and mood associated with the performance of the task.

METHODS

The study was reviewed and approved by the University of Nottingham Medical School Research Ethics Committee (Approval Reference; L13092012 SCS Anesthesia). Medical students from the University of Nottingham Medical School were invited to participate in the study through poster advertising. The poster [appendix 1] was distributed via e-mail, and displayed on the Networked Learning Environment (NLE), by and with the permission of the medical student undergraduate co-ordinator (University of Nottingham). Students who expressed a wish to participate were emailed a participant information leaflet [appendix 2-5] and an invitation to attend the study. Written informed consent [appendix 6] for the study, including video recording, was taken from all participants.

Design:

This single center, prospective, blinded observational study was conducted at the University Department of Anesthesia, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham UK.

Subjects with previous experience of ultrasound scanning or of performing regional anesthesia were excluded from the study. The study was organized in four phases [figure 2]. The enrolled medical students were asked to undergo and complete all four phases of the study. Participants' identities were masked throughout the study, and their assessments were concealed from view within individual folders. Assessors of the ultrasound-guided needling task (phase four) were blinded to the outcomes of the preceding assessments.
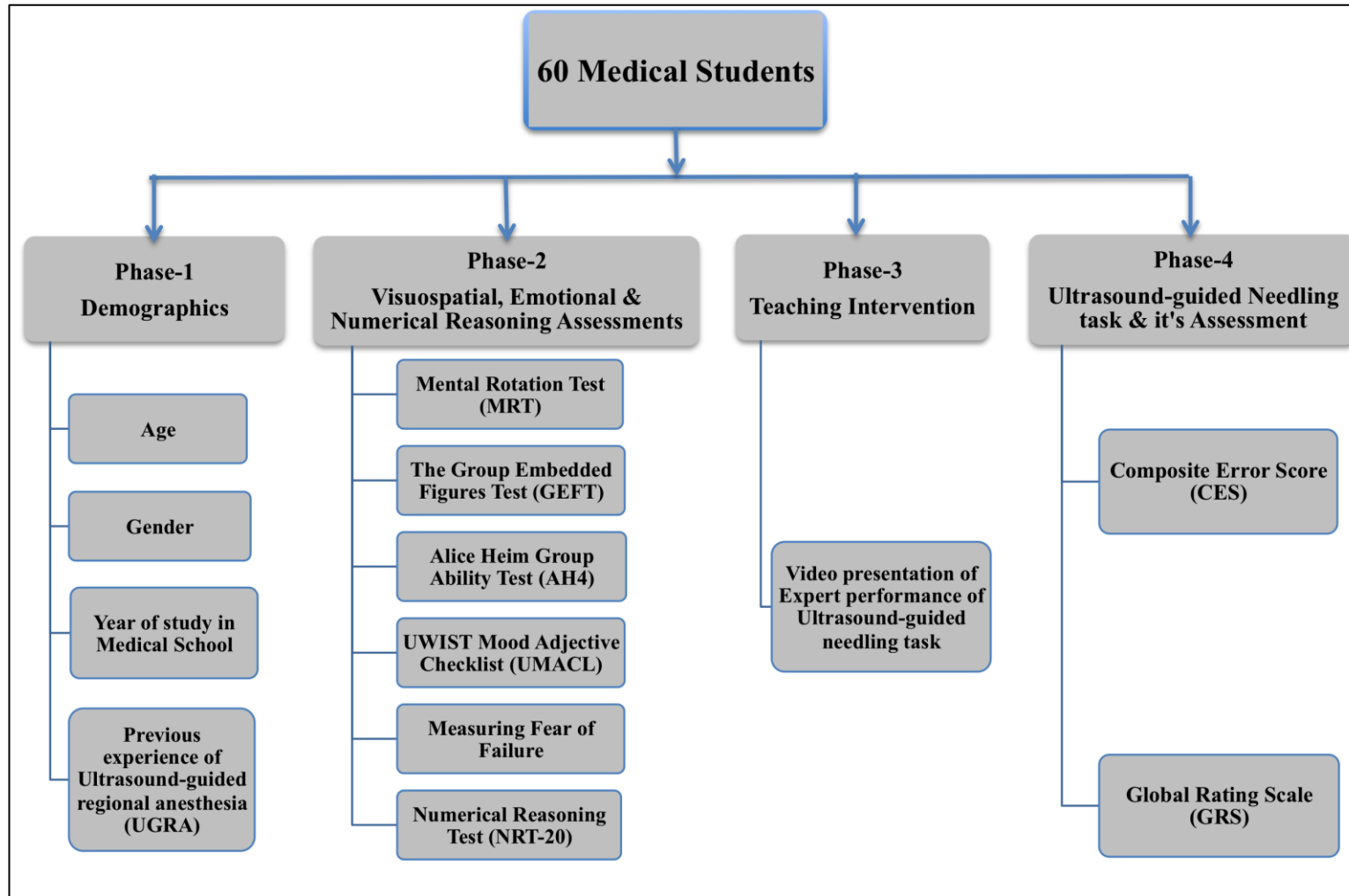
Figure 2 Flowchart showing study design

PHASE ONE

This phase aimed at collecting basic demographic data including age, sex, year of study in medical school and previous experience of UGRA.

PHASE TWO

This phase consisted of visuospatial, emotional, and numerical reasoning assessments of the study participants. The assessments were paper-based and administered under examination conditions. Participants were blinded to the study hypothesis and their test scores. Brief descriptions of all the visuospatial, emotional, and numerical reasoning assessments are provided below.

Visuospatial Assessments:

The visuospatial assessments consisted of the Mental Rotation Test (MRT), Group Embedded Figures Test (GEFT) and Alice Heim Group Ability Test (AH4).

Mental rotation Test (MRT)[180-182]

There are four different variations, based on the original Vandenberg & Kuse (1978) mental rotation test figures, which in turn are based on figures provided by Shepard & Metzler.[183] These include MRT–A, MRT–B, MRT–C and MRT–D respectively.

MRT – A: This is the standard set, with stimulus figures redrawn from the original Vandenberg & Kuse set.

MRT – B: This consists of the same items as MRT – A, but reshuffled in a different order. This test is meant to be used as alternative test of similar difficulty as MRT – A. In practice, it seems to have consistently lower means than MRT – A, but the difference is not significant.

MRT – C: This is the most difficult version of the mental rotation test. It requires the subject to rotate the stimuli in two directions. While MRT – A and MRT – B engages mental rotation around the vertical axis, this test involves mental rotation around both the vertical and horizontal axis, respectively.

MRT – D: This test is similar to MRT – A, with the exception that all stimulus and target figures are rotated 90 degrees to the left. In order to solve the problems, subjects have to rotate the figures around the horizontal axis. This appears to be more difficult than rotation around the vertical axis.

There are two forms of the original Vandenberg & Kuse test in the literature, one with 20 and one with 24 problem sets. We used MRT–A, which consists of 24 problem figures. Each problem task has a target figure on the left and four stimulus figures on the right. Two of these stimulus figures are rotated versions of the target figure and two of the stimulus figures cannot be matched to the target figure. The aim is to mentally rotate the figures around the vertical axis to find the two correct rotated versions of the target figure. Participants were given four minutes to complete the first set of 12 problem tasks, followed by a one-minute break before completing the second set of 12 problem tasks in the next four minutes.

Two ways of scoring are found in the literature. The first gives one point for each correct answer, and a point is subtracted for each incorrect answer. This would yield a maximum of 48 points. However, we used the other method of scoring, where one and only one point is given if both of the stimulus figures that match the target figure are identified correctly. No point is given for a single correct answer. This means that the maximum score obtainable is 24 points.

The Group Embedded Figures Test (GEFT)[184,185]

This test was developed by Philip K. Oltman, Evelyn Raskin, and Herman A. Witkin. The GEFT was designed to provide an adaptation of the original individually administered Embedded Figures Test (EFT), which would make possible group testing. The use of the individually administered EFT is often impractical where large numbers of subjects must be tested for screening on the field-dependence-independence dimension. The GEFT has been modeled as closely as possible on the individually administered EFT with respect to mode of presentation and format. It contains 18 complex figures, 17 out of which were taken from the EFT. The GEFT measures field-independence, which is the ability to perform a focal task independently of any background information or distracters. The aim is to find a previously seen simple figure within a larger complex figure, which has been structured in a way to obscure or embed the simple figure. The participants were required to identify and outline accurately a simple shape embedded in a complex figure. The test consists of three sections. The participants were initially given two minutes to complete the seven problems in the first section. Following this, second and third sections consisting of nine problems each respectively were completed in 10 minutes.

The score is the total number of simple forms correctly traced in second and third sections combined. Omitted items are scored as incorrect. The items in the first section are not included in the total score. However, the scorers scanned this section before starting to score the test as a means of making sure that the subject has understood the test directions. In order to receive the credit for an item, all lines of the simple form must be traced; the subject has added no extra lines and all incorrect lines have been erased. In this way the maximum score a subject could attain is 18.

Alice Heim Group Ability Test (AH4)[186]

The AH4 is designed as a group test of general intelligence, which primarily assesses deductive reasoning including verbal, mathematical and spatial reasoning. Many current intelligence tests consist of problems involving only one 'bias' and only one type of principle. The aim in AH4 is to incorporate as many different 'biases' and principles as is consistent with a reasonably short test and with the inclusion of preliminary examples, which are illustrative of all the principles, found in the test proper. AH4 incorporate 2 parts. Part-1 consists of 65 questions that have a verbal or numerical bias. Six types of principle are included: directions, verbal opposites, numerical series, verbal analogies, simple arithmetical computations and synonyms. Part-2 comprises of 65 questions which have a diagrammatic bias and which exemplify five types of principle: analogies, sames, subtractions, series and superimpositions. These are in multiple-choice form, the number of proffered answers in every case being five. In both parts of the test, the principles are presented cyclically, in the orders given above and the

problems are arranged roughly in ascending order of difficulty within each principle.

In our study we used AH4 to only assess diagrammatic bias in the form of spatial reasoning skills, which is the ability to visualize, mentally rotate and manipulate two-dimensional or three-dimensional shapes or patterns. As mentioned above the test consists of 65 questions and the time limit for completing the test is 10 minutes, exclusive of the preliminary examples. One mark is scored for each correct answer. Therefore the maximum score that could be achieved is 65.

Emotional Assessments:

Emotional processes that tap into state anxiety or tense arousal, as well as positive mood states (e.g., energetic arousal), were assessed using the UWIST Mood Adjective Checklist (UMACL) and Fear of Failure.

UWIST Mood Adjective Checklist (UMACL)[187]

The UMACL is used to measure mood and comprises three bipolar scales: energetic arousal (EA) [vigorous vs. tired: coefficient alpha = .79], tense arousal (TA) [nervous vs. relaxed: coefficient alpha = .76] and hedonic tone (HT) [pleasant vs. unpleasant mood: coefficient alpha = .81].[188] In addition to these scales, a mono-polar anger/frustration (AF: coefficient alpha = .80) scale was also used. The participants were instructed to complete the UMACL checklist according to their present mood using 28 adjectives each on a four – point scale ('definitely', 'slightly', 'slightly not' or 'definitely not'). Scores were obtained by

summating item scores, but on the UMACL, positive items are 'reverse-scored'. Therefore subtracting each positive item from '5' recalculated the positive items, before item scores are added together.

Fear of Failure[189]

'Fear of failure' assesses the general preference to be motivated not to succeed but to avoid failing. This assessment consists of four statements pertaining to fear of failure, each scored on a four – point scale ('always', 'often', 'rarely' and 'never'). The participants had to mark one option, which is most relevant to them. Scores were then obtained by summating the item scores. The reported coefficient alpha was .70.

Numerical Reasoning Assessments:

Numerical Reasoning Test (NRT-20)[190]

This test measures mathematical and logical reasoning via 20 short reasoning problems based on numbers that do not require any previous training in mathematics. It is a test of fluid intelligence, which depicts skills of problem solving, abstract reasoning, and ability to learn new things, irrespective of prior knowledge or education. There are 20 items, which include series completion (numbers and matrices), basic arithmetic problems (computational speed), and other deductive reasoning tasks. The participants were given 15 minutes to solve as many problems as possible.

PHASE THREE

The third phase is a teaching intervention. In this phase the participants were given 30 minutes to watch and review an 11 – minute video[191] presentation mapped to specific learning objectives, which modeled expert performance of ultrasound-guided needle advancement in a turkey breast model.

The learning objectives were:

I. Switch on the ultrasound machine (S-Series, Sonosite Limited, Hitchen, UK).

II. Correctly orientate the ultrasound probe (linear, 38mm) in relation to the display on the screen.

III. Ensure adequate application of conducting gel to enhance ultrasound transmission and picture quality.

IV. Locate and identify the target (olive) within the turkey breast.

V. Adjust the gain function to improve the quality of the image by altering brightness of the picture.

VI. Alter the depth of the image to obtain a suitable image of the target.

VII. Using an in-plane approach, insert a 50mm Stimuplex® A needle (B. Braun, Melsungen, Germany) into the turkey breast and aim to place the needle tip at the 12 o'clock position, as indicated by the attending assessors, above the upper edge of the target, without piercing the target.

PHASE FOUR

The fourth phase included an ultrasound-guided needling task and its assessment. Participants were asked to complete the ultrasound-guided needling task, as demonstrated in the video, on a turkey breast model[192,193] using a standard ultrasound transducer probe (38-mm high-frequency linear array transducer; HFL38X 13-6 MHz, Sonosite Limited, Hitchen, UK). In order to improve realism, the turkey breast model was inserted into the draped groin recess of a Laerdal® IV Torso manikin (Laerdal Medical Limited, Orpington, Kent, UK), which was used solely for this study. Each participant's performance of the ultrasound-guided needling task was video recorded on a tripod-mounted camcorder focusing the whole task including the ultrasound image. The rationale of video recording was for reference purposes, in case if there were any uncertainties concerning the participants' performance. The participants received no help or feedback before or during the task. Study participation ceased once the ultrasound-guided needling task was completed.

Two anesthesiologists experienced in UGRA independently assessed the participants' performance of the task by means of two previously validated assessment tools; composite error score (CES)[127,193,194] [appendix 7] and global rating scale (GRS)[106,117,146] [appendix 8]. The assessors had undergone specific training and practice in the use of these assessment tools. The CES used in our study was similar to the one utilized by Davies T et al.[194] The CES was calculated by adding together the measurements of 'total number of errors', 'number of needle passes' and 'image quality scale' for each of the participant's task performance. 'Number of needle passes' is defined as, *"a new puncture of the turkey breast skin or the withdrawal of the needle towards the exterior of the turkey breast skin"*. This definition along with

the classification of 'image quality scale' is similar to the one used by Sites BD et al.[193]

A lower composite error score is associated with better accuracy and task performance. The GRS was a modified version of that used in the original study by Martin JA et al.[120] The GRS consisted of seven items each rated on a five – point scale. The GRS predominantly assessed more general behaviors and the overall performance of the participant. For analysis purposes, the CES and GRS were deconstructed into their 'type of errors and performances' respectively. Each type of error and performance was then correlated with a specific visuospatial assessment that was believed to be generally inter-related, according to the expert Psychologist.

Outcomes measured:

Primary outcomes:

I. Correlations between CES and visuospatial assessments score

II. Correlations between GRS and visuospatial assessments score

Secondary outcomes:

I. Reliability of the CES & GRS

II. Correlation between CES & GRS

We examined the following relationships:

1. *CES & type of errors with visuospatial, emotional and numerical reasoning assessments:*

   a) The MRT and AH4 will be separately correlated with the following;

      I. CES

      II. Advancing the needle without visualising the needle tip

      III. Failure to recognise the orientation of the ultrasound probe with the image on the screen

      IV. Malpositioning of target on screen including incorrect selection of depth of image

      V. The number of needle passes

   b) The GEFT will be correlated with the following;

      I. CES

      II. Failure to identify the target

   c) UMACL, fear of failure and NRT-20 will be correlated with the following;

      I. CES

      II. Unintentional probe movement

      III. Focussing on hand position rather than ultrasound image, as needle tip approaches target

2. *GRS & type of performances with visuospatial, emotional and numerical reasoning assessments:*

   a) The MRT and AH4 will be separately correlated with the following;

      I.   GRS

      II. Time & motion

      III. Instrument handling

      IV. Flow of procedure

   b) The GEFT will be correlated with;

      I.   GRS

      II. Image quality

   c) UMACL, fear of failure and NRT-20 will be correlated with the following;

      I.   GRS

      II. Preparation for procedure

      III. Flow of procedure

      IV. Image quality

Statistical Analysis:

Descriptive statistics for demographic and outcome measure data were calculated. CES data follow a non-normal distribution, and are thus presented as median (IQR). Normality of other data was assessed by histogram and the Shapiro-Wilk and Skewness / Kurtosis tests.

We performed an initial exploratory analysis using Spearman's correlation coefficient 'ρ' (Rho) to determine which of the six explanatory variables was the most predictive of better task performance. The count data (CES) were over-dispersed. This was unlikely due to excessive zeros as the proportion of extra zeros was considerably small (4/60 = 6.66 %); therefore negative-binomial regression analysis was conducted for CES. For continuous data the relationship between each potential explanatory variable and GRS was evaluated in an ordinary least square (OLS) regression. Bonferroni corrections were applied for multiple testing. We then created a regression model using the explanatory variable most predictive of performance. To achieve a study power of 0.8 ($\alpha = 0.05$), we calculated that we would need to recruit 60 participants for this model with an assumed moderate effect size[195] ($r = 0.3 - 0.5$).

We chose to make non-pairwise comparisons between males and females in order to determine whether any differences in visuospatial ability existed. In all cases, we used P – values less than 0.05 (two-tailed) to indicate statistical significance.

Reliability of the assessment tools i.e. CES and GRS, was evaluated using Intra-class correlation coefficient (ICC), Cronbach's alpha coefficient and standard error of the measurement as a percentage of the mean (SEM%).[196,197] The statistical analysis software STATA/IC version 10.0 (StataCorp, Texas, USA) was used for data analysis.

RESULTS

Participant demographics and summary statistics for visuospatial ability and task performance are summarized in table 1. Males were found to exhibit better mental rotation skills compared to females (P < 0.001) [table 2].

Table 1 Participant Demographics and Summary Statistics for Visuospatial ability and Task performance

Participant Demographics	
Characteristics	**Overall N=60**
Age in years, Mean (SD)	23.3 (4.3)
Gender	
Male (%)	30 (50)
Female (n, %)	30 (50)
Year of study in medical school*	
Year 1 (%)	08 (14.0)
Year 2 (%)	21 (36.8)
Year 3 (%)	13 (22.8)
Year 4 (%)	11 (19.3)
Year 5 (%)	04 (7.0)

Summary Statistics	
Assessment	**Score**
MRT (mean, SD)	13.9 (5.3)
GEFT (median, IQR)	17.0 (14.0 – 17.5)
AH4 (median, IQR)	60.0 (53.0 – 63.0)
[UMACL] EA (median, IQR)	19.0 (16.5 – 21.5)
[UMACL] TA (mean, SD)	15.2 (3.4)
[UMACL] HT (median, IQR)	27.0 (25.0 – 29.0)
[UMACL] AF (median, IQR)	6.0 (5.0 – 9.0)
Fear of failure (mean, SD)	6.8 (1.6)
NRT-20 (mean, SD)	13.9 (2.5)
CES (median, IQR)	6.0 (3.0 – 10.0)
GRS (mean, SD)	19.7 (6.0)

* 3 (5%) missing
MRT – Mental Rotation Test; GEFT – The Group Embedded Figures Test; AH4 – Alice Heim Group Ability Test; [UMACL] – UWIST Mood Adjective Checklist; EA – Energetic Arousal; TA – Tense Arousal; HT – Hedonic Tone; AF – Anger/Frustration; NRT-20 – Numerical Reasoning Test; CES – Composite Error Score; GRS – Global Rating Scale.

Table 2 MRT scores according to sex of participants

Gender	Number (%)	MRT (Mean)	SD	95 % CI	P Value
Female	30 (50)	10.9	4.64	9.16 – 12.63	
					<0.001
Male	30 (50)	17.0	4.15	15.44 – 18.55	

Values are Mean and SD – Standard deviation; CI – Confidence Interval; MRT – Mental Rotation Test.

Reliability of CES and GRS:

The intra-class correlation coefficient and SEM (%) for CES and GRS was 0.97 (15.29%) and 0.91 (8.53%) respectively; this demonstrates a high degree of inter-rater agreement. Similarly, Cronbach's alpha coefficient and SEM (%) for CES and GRS was 0.98 (9.49%) and 0.96 (5.69%) respectively; this demonstrates a high degree of inter-item consistency.

Composite error score (CES) versus visuospatial ability assessments:

Of the three visuospatial assessments (MRT, GEFT, AH4), only MRT correlated significantly with the CES ($\rho = -0.54$; $P < 0.001$) [figure 3, table 3], indicating that a low error rate is associated with high MRT scores or vice versa. The negative binomial regression coefficients for each of the variable showed that for each unit increase in MRT, the expected log count of the CES decreases by 0.08-unit ($P < 0.001$).
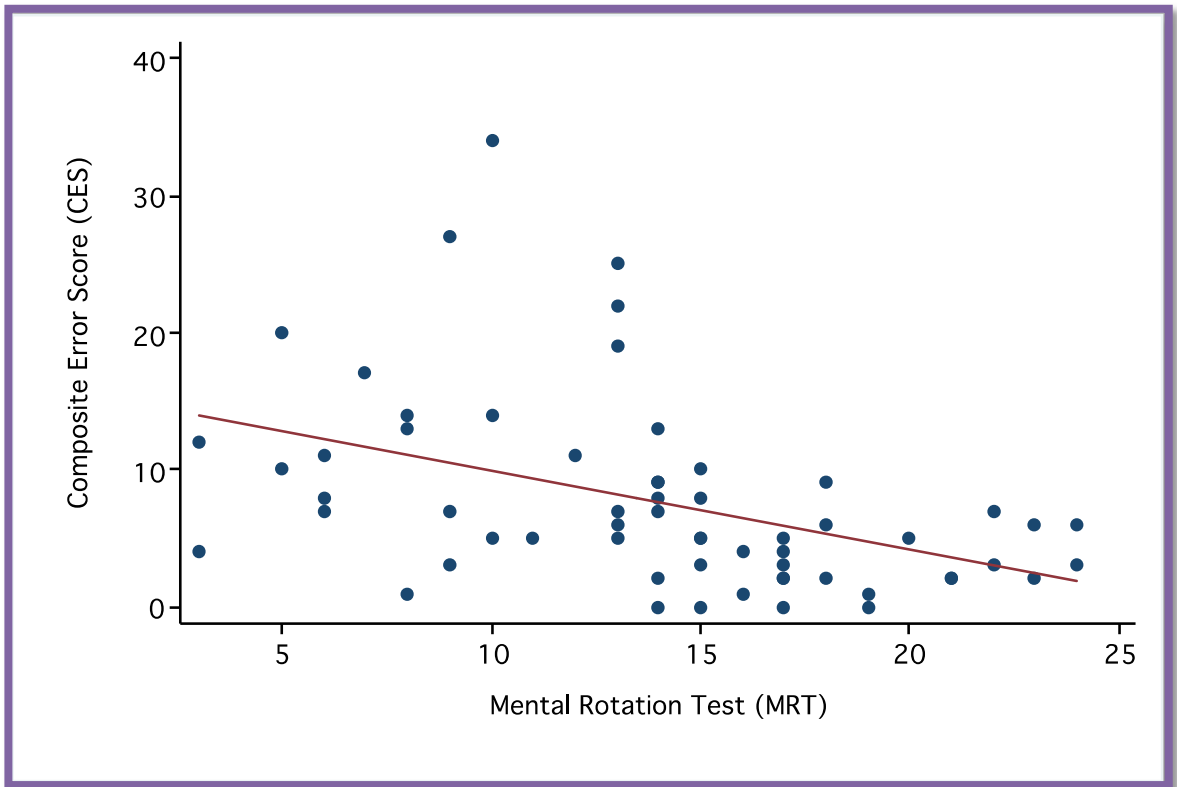
Figure 3 Relationship of CES with MRT

MRT is negatively correlated with CES ($\rho = -0.54$; $P < 0.001$), which reveals that increasing error rate is associated with low MRT scores.

Table 3 Correlation of CES & GRS with Visuospatial, Emotional and Numerical Reasoning Assessments

Visuospatial, Emotional & Numerical Reasoning Assessments	Spearman's correlation coefficient (ρ)	
	CES	GRS
MRT	−0.54 (P < 0.001)	0.47 (P < 0.001)
AH4	−0.09 (P = 0.49)	0.09 (P = 0.49)
GEFT	−0.06 (P = 0.60)	0.00 (P = 0.95)
UMACL		
• *Energetic Arousal (EA)*	−0.30 (P = 0.01)	0.06 (P = 0.62)
• *Tense Arousal (TA)*	0.26 (P = 0.04)	−0.35 (P = 0.005)
• *Hedonic Tone (HT)*	−0.22 (P = 0.08)	0.18 (P = 0.15)
• *Anger/Frustration (AF)*	0.16 (P = 0.21)	−0.19 (P = 0.13)
Fear of Failure	0.05 (P = 0.69)	−0.07 (P = 0.57)
NRT-20	0.01 (P = 0.93)	−0.05 (P = 0.69)

Spearman's correlation coefficient (ρ) of Composite Error Score (CES) and Global Rating Scale (GRS) with visuospatial, emotional and numerical reasoning assessments; Significance at P < 0.05. MRT – Mental Rotation Test; AH4 – Alice Heim Group Ability Test; GEFT – The Group Embedded Figures Test; [UMACL] – UWIST Mood Adjective Checklist; NRT-20 – Numerical Reasoning Test.

On the other hand, GEFT ($\rho = -0.06$; P = 0.60) [table 3] and AH4 ($\rho = -0.09$; P = 0.49) [table 3] did not demonstrate a significant relationship with the CES.

The relationships between each type of error and visuospatial ability assessments are presented in table 4. After Bonferroni adjustments, two errors described as *'needle advancement without visualization of needle tip'* ($\rho = -0.52$) and *'number of needle passes'* ($\rho = -0.45$) correlated significantly with MRT (both, P < 0.0016) [table 4].

Table 4 Correlation of type of Errors with Visuospatial, Emotional and Numerical Reasoning Assessments

Type of Errors	Visuospatial, Emotional & Numerical Reasoning Assessments	
	MRT	**AH4**
1- Needle advanced without visualisation of needle tip	−.52 (P < 0.001)	−.21 (P = 0.09)
2- Failure to recognise orientation of probe with the image screen	−.24 (P = 0.06)	−.10 (P = 0.43)
3- Target malpositioned on screen, including incorrect depth selection	−.08 (P = 0.50)	−.01 (P = 0.93)
4- Number of needle passes	−.45 (P < 0.001)	−.11 (P = 0.36)
	GEFT	
1- Failure to identify the target	.06 (P = 0.61)	
	UMACL: (EA, TA, HT, AF)	**Fear of failure** / **NRT-20**

Type of Errors	UMACL: (EA, TA, HT, AF)	Fear of failure	NRT-20
1- Unintentional probe movement	−.38 (P = 0.002) EA .13 (P = 0.29) TA −.17 (P = 0.18) HT .07 (P = 0.58) AF	.00 (P = 0.98)	.16 (P = 0.21)
2- Attention focussed on hand and not image as needle advanced	−.17 (P = 0.18) EA −.01 (P = 0.93) TA .08 (P = 0.51) HT −.00 (P = 0.98) AF	.06 (P = 0.62)	.27 (P = 0.03)

Spearman's correlation coefficient (Rho) was used in the exploratory analysis to determine which of the six explanatory variables were associated with type of errors.
For type of errors, the Bonferroni correction applied for multiple comparisons and P value was set at < 0.0016.
MRT – Mental Rotation Test; AH4 – Alice Heim Group Ability Test; GEFT – The Group Embedded Figures Test; [UMACL] – UWIST Mood Adjective Checklist; EA – Energetic Arousal; TA – Tense Arousal; HT – Hedonic Tone; AF – Anger/Frustration; NRT-20 – Numerical Reasoning Test; CES – Composite Error Score.

Global rating scale (GRS) versus visuospatial ability assessments:

Out of the three visuospatial assessments (MRT, GEFT, AH4), only MRT correlated significantly with the GRS ($\rho = 0.47$; $P < 0.001$) [figure 4, table 3], indicating that better performance was associated with better mental rotation skills or vice versa. An ordinary least square regression established the univariate association of GRS with MRT showing that for a one-unit increase in MRT we would expect a 0.43-unit increase in GRS ($P = 0.002$).
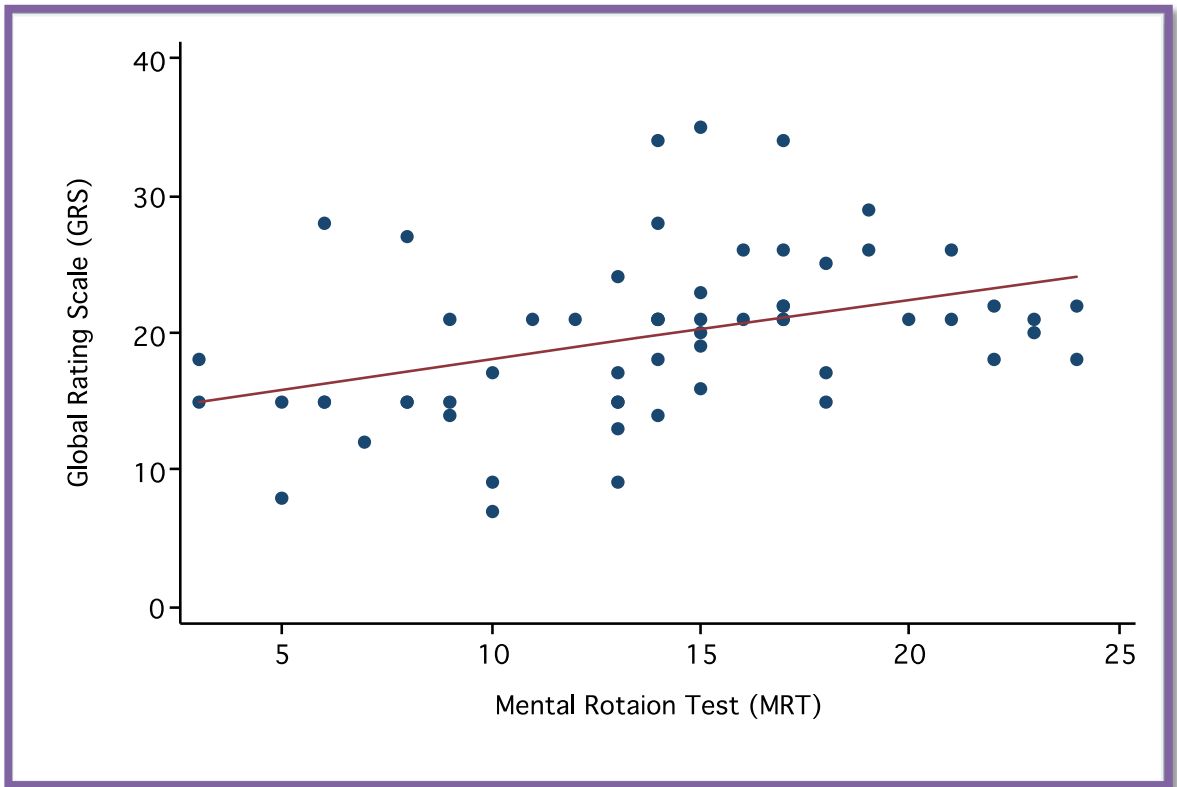
Figure 4 Relationship of GRS with MRT

MRT is positively correlated with GRS ($\rho = 0.47$; $p < 0.001$), which shows that enhanced GRS quality scores are associated with high MRT scores.

On the other hand, GEFT ($\rho = 0.00$; P = 0.95) [table 3] and AH4 ($\rho = 0.09$; P = 0.49) [table 3] did not demonstrate a significant relationship with the GRS.

The relationships between each domain of the GRS and visuospatial ability assessments are presented in table 5. After Bonferroni adjustments *'time and motion'* ($\rho = 0.44$), *'instrument handling'* ($\rho = 0.47$) and *'flow of procedure'* ($\rho = 0.44$) correlated significantly with MRT (all three, P < 0.0031) [table 5].

Table 5 Correlation of Type of Performances with Visuospatial, Emotional and Numerical Reasoning Assessments

Type of Performances	Visuospatial, Emotional & Numerical Reasoning Assessments		
	MRT		**AH4**
1- Time and motion	.44 (P < 0.001)		.14 (P = 0.25)
2- Instrument handling	.47 (P < 0.001)		.07 (P = 0.55)
3- Flow of procedure	.44 (P < 0.001)		.06 (P = 0.59)
	GEFT		
1- Image quality	−.00 (P = 0.96)		
	UMACL: (EA, TA, HT, AF)	**Fear of failure**	**NRT-20**
1- Preparation for procedure	.02 (P = 0.87) EA −.28 (P = 0.02) TA .02 (P = 0.86) HT −.07 (P = 0.58) AF	−.00 (P = 0.97)	−.01 (P = 0.89)
2- Flow of procedure	.07 (P = 0.58) EA −.30 (P = 0.01) TA .19 (P = 0.13) HT −.10 (P = 0.42) AF	−.02 (P = 0.86)	−.03 (P = 0.76)
3- Image quality	.21 (P = 0.10) EA −.28 (P = 0.02) TA .14 (P = 0.28) HT −.23 (P = 0.06) AF	−.11 (P = 0.38)	−.05 (P = 0.66)

Spearman's correlation coefficient (Rho) was used in the exploratory analysis to determine which of the six explanatory variables were associated with type of performances.
For type of performances, the Bonferroni correction applied for multiple comparisons and P value was set at < 0.0031.
MRT – Mental Rotation Test; AH4 – Alice Heim Group Ability Test; GEFT – The Group Embedded Figures Test; [UMACL] – UWIST Mood Adjective Checklist; EA – Energetic Arousal; TA – Tense Arousal; HT – Hedonic Tone; AF – Anger/Frustration; NRT-20 – Numerical Reasoning Test; GRS – Global Rating Scale.

Composite error score (CES) versus emotional and numerical reasoning assessments:

Of the UMACL, energetic arousal (EA) was found to correlate significantly with the CES ($\rho = -0.30$; $P = 0.01$) [figure 5, table 3]. By contrast, tense arousal (TA) correlated weakly with the CES ($\rho = 0.26$; $P = 0.04$) [figure 6, table 3]. This showed that errors would be low in vigorous, active individuals and high in anxious individuals. The negative binomial regression coefficients for each of the variable showed that for each unit increase on EA, the expected log count of the CES decreases by 0.07-unit ($P = 0.02$).
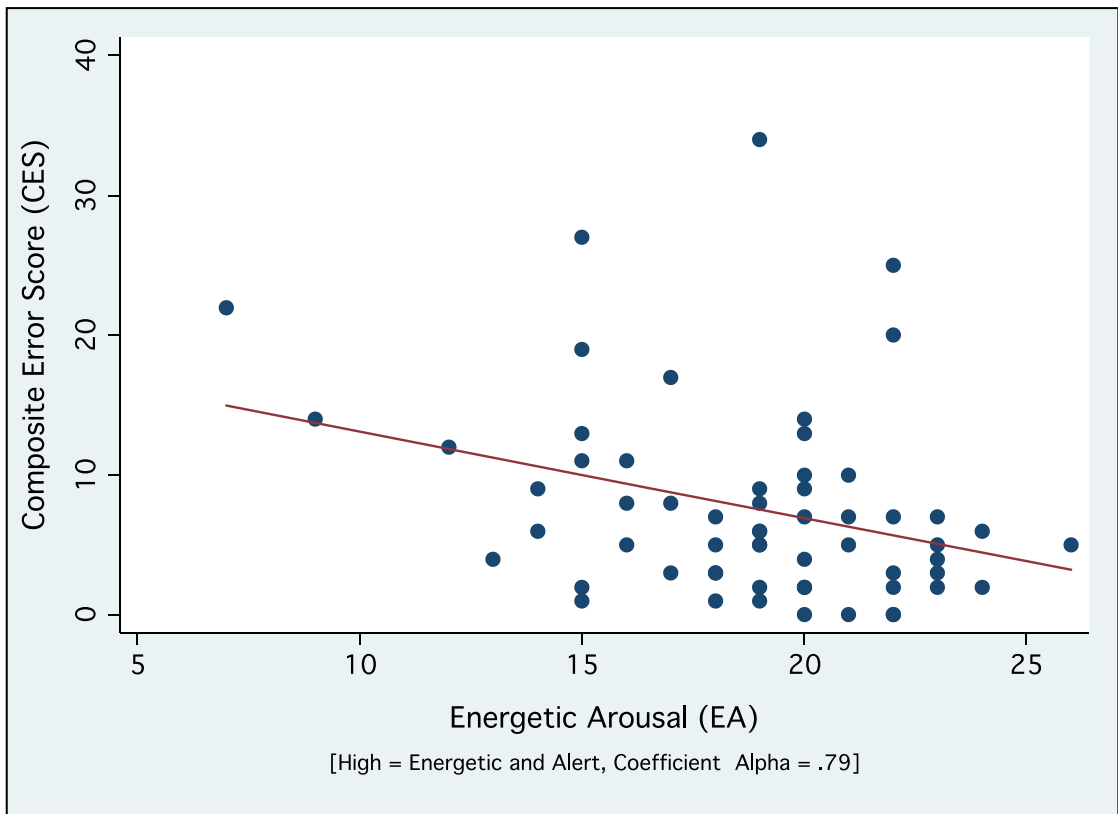
Figure 5 Relationship of EA with CES

EA is negatively correlated with CES ($\rho = -0.30; P = 0.01$), which demonstrates that low error rates are associated with energetic and active individuals.
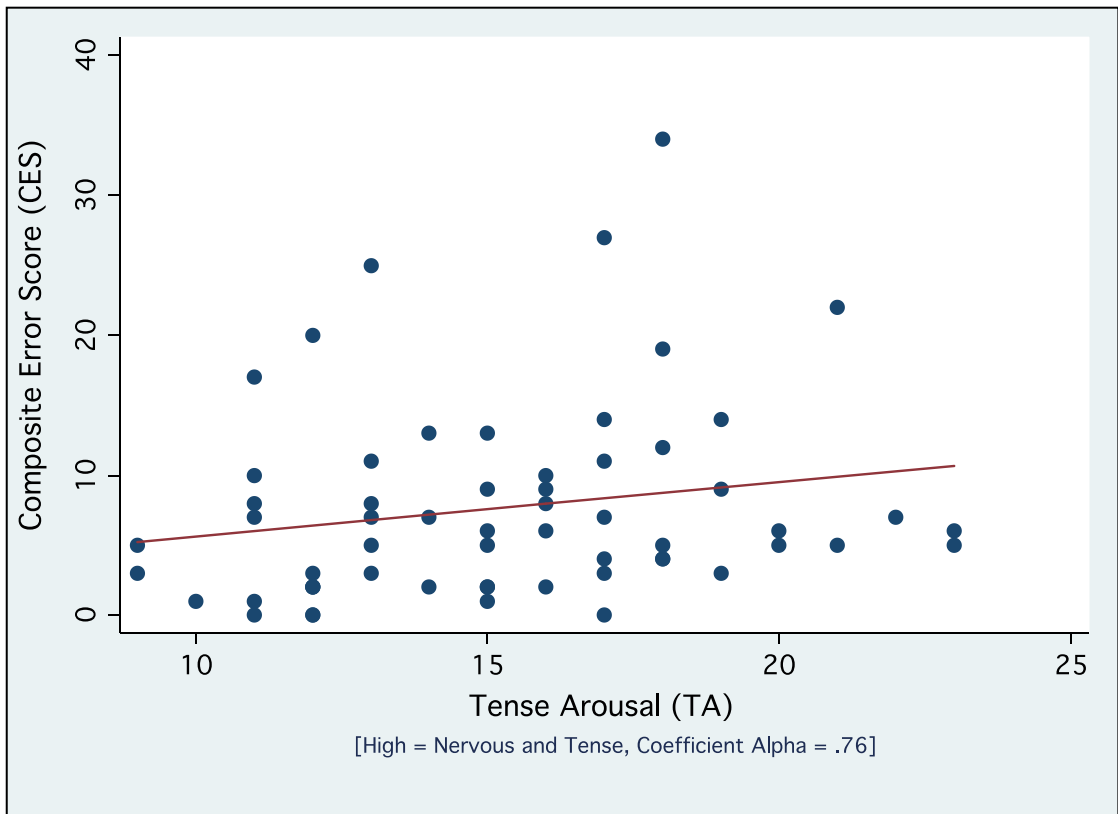
Figure 6 Relationship of TA with CES

TA is positively correlated with CES ($\rho = 0.26$; $P = 0.04$), which illustrates that high error rates are associated with tense, anxious and nervous individuals.

Among other sub-types of UMACL, hedonic tone (HT) ($\rho = -0.22$; P = 0.08) [table 3] and anger/ frustration (AF) ($\rho = 0.16$; P = 0.21) [table 3] did not demonstrate a significant relationship with the CES. Similarly, fear of failure ($\rho = 0.05$; P = 0.69) [table 3] and NRT-20 ($\rho = 0.01$; P = 0.93) [table 3] did not reveal any significant relationship with the CES.

The relationships between each type of error and emotional, and numerical reasoning assessments are presented in table 4. After Bonferroni adjustments, type of error described as *'unintentional probe movement'* and *'attention focussed on hand and not image as needle advanced',* did not correlate significantly with UMACL, fear of failure and NRT-20, respectively [table 4].

Global rating scale (GRS) versus emotional and numerical reasoning assessments:

Of the UMACL, only tense arousal (TA) correlated significantly with the GRS ($\rho = -0.35$; P = 0.005) [figure7, table 3]. This showed that GRS quality scores would be low when anxiety level is high or vice versa. An ordinary least square regression established the univariate association of GRS with TA showing that for a one-unit increase in TA we would expect a 0.54-unit decrease in GRS (P = 0.01).
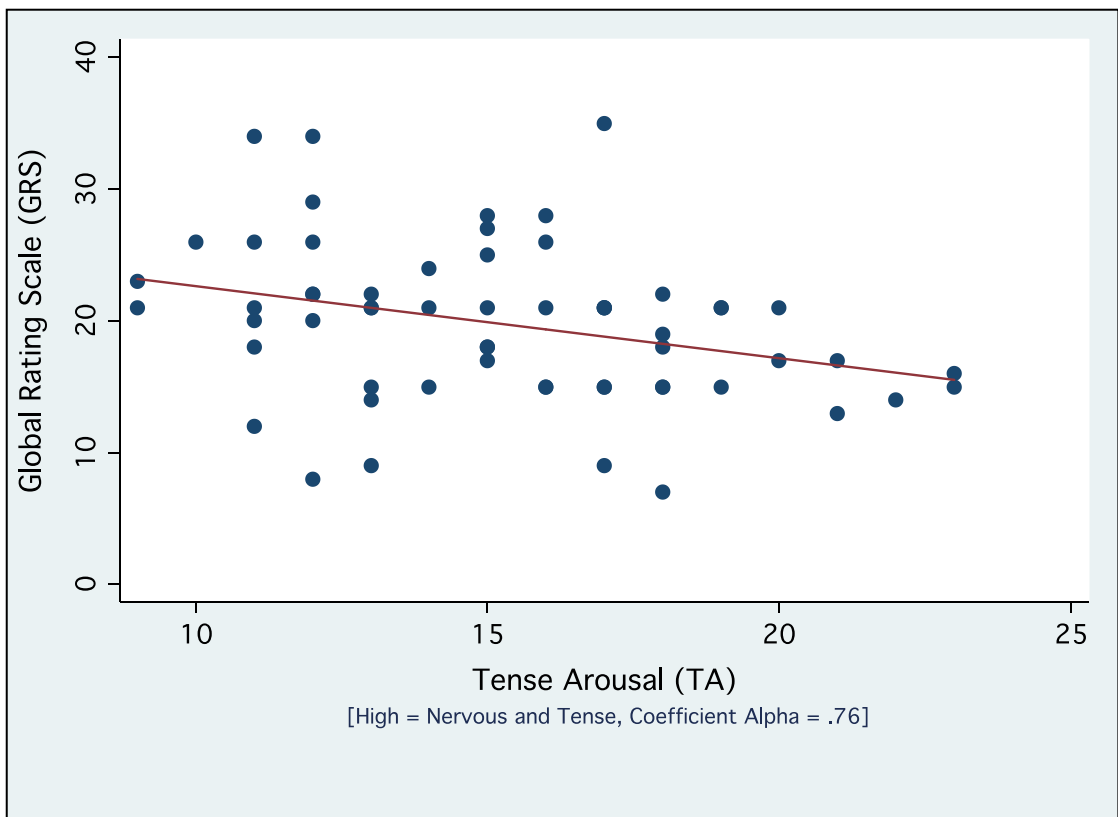
Figure 7 Relationship of TA with GRS

TA is negatively correlated with GRS ($\rho = -0.35$; P = 0.005), which explains that low GRS quality scores are associated with tense, anxious and nervous individuals.

Including other sub-types of UMACL, energetic arousal (EA) (ρ = 0.06; P = 0.62) [table 3], hedonic tone (HT) (ρ = 0.18; P = 0.15) [table 3] and anger/ frustration (AF) (ρ = $-$ 0.19; P = 0.13) [table 3] did not demonstrate a significant relationship with the CES. Similarly, fear of failure (ρ = $-$ 0.07; P = 0.57) [table 3] and NRT-20 (ρ = $-$ 0.05; P = 0.69) [table 3] did not reveal any significant relationship with the GRS.

The relationships between each type of performance and emotional, and numerical reasoning assessments are presented in table 5. After Bonferroni adjustments, type of performance described as *'preparation for procedure', 'flow of procedure'* and *'image quality',* did not correlate significantly with UMACL, fear of failure and NRT-20, respectively [table 5].

DISCUSSION:

This study signifies the ability of visuospatial, emotional and numerical reasoning skills in predicting the novice performance of UGRA. Among the visuospatial skills, we found that MRT correlated positively with the GRS and negatively with the CES. This would mean that participants with better MRT scores did fewer errors and came up with good performance scores when completed an ultrasound-guided skill task, thus demonstrating an overall good performance. MRT has been previously reported to correlate positively with the performance in simulated laparoscopic studies.[159,160,166,168,198] However, it is unique to this study that MRT has been applied for the first time in the performance of UGRA. Other visuospatial assessments including the GEFT and AH4 instinctively anticipated to yield similar results but failed to illustrate any association with the CES and GRS respectively. The results indicate that MRT predicts novice performance of an ultrasound-guided needle advancement task in a turkey breast model. As a trait measure, MRT has the potential to be used as a tool to focus educational and training resources on individuals who have less ability to perform ultrasound-guided needle tasks.

The study shows that males are likely to have better mental rotation capabilities than females. This is in line with two previous meta-analyses[199,182] which showed effect sizes around 0.95 favoring males. The results of the meta-analysis done by Linn MC and Petersen AC[199] suggest that sex differences arise on some types of spatial ability but not others, and large sex differences are found only on measures of mental rotation, while smaller sex differences are found on measures of spatial perception. This difference in male and female means that the MRT cannot be used as a selection

tool for medical posts, i.e. high stakes assessment, as men are likely to be preferentially selected and this would introduce indirect sexual discrimination against women. However, that is not to say that males with better MRT scores perform better at the task. Two previous studies of laparoscopic skills have demonstrated that gender does not affect psychomotor performance[200,201], though the affect of MRT scores of both males and females remained unknown. Kolozsvari NO et al.[200] evaluated the impact of gender on the learning curve of medical students for a fundamental laparoscopic task. They found that there were no differences between men and women in initial peg transfer score, learning plateau, or learning rate and gender had no significant relationship to early performance. Similarly, in another observational study of surgical trainees with very limited laparoscopic experience, Grantcharov TP et al.[201] demonstrated that male trainee surgeons made a similar number of errors and unnecessary hand movements as females during their performance of six simulated laparoscopic tasks.

One may suggest that it would be useful to provide a range of MRT scores wherein learners could benefit the most from focused training. However, we are unable to do so at this stage of our work. While our study shows that MRT has a predictive validity for performance of an ultrasound-guided needle task, it does not indicate at what point MRT performance can be defined as adequate or not. Despite this limitation, we believe that it is reasonable to state that low error rates, better image quality, and better global performance are associated with higher MRT scores. By implication strategies, which aim to develop mental rotation skills, may be developed and employed in an attempt to improve ultrasound-guided needle advancement skills.

Out of the emotional assessments, the UMACL and its sub-type of mood called "Tense Arousal (TA)" correlated positively (but weakly) with the CES and negatively with the GRS. That is, tensed and nervous participants did more errors and performed less well in the ultrasound-guided needling task. On the other hand, UMACL's another sub-type of mood termed as "Energetic Arousal (EA)", correlated negatively with the CES, indicating that energetically aroused and vigorous participants did fewer errors.

Hence we found that negative mood adversely affects performance. Good performance, therefore, appears to be a function of visuospatial ability and reducing anxiety. This is particularly important in avoiding all the potential sources that could build up inner tension and result in an overall poor performance. Thus, a second practical implication of our study is that stress and anxiety when performing the task may need to be reduced in the training and learning environment, and therefore interventions to achieve this need to be developed. However, in the real clinical setting some level of anxiety will be associated with any clinical intervention, and thus the degree of stress elicited by the task may itself increase its validity.[202] Thus, reducing stress in the training context may not be warranted. By extension, it may be more relevant for educators to develop curricula that allow novice practitioners to learn the necessary coping skills to deal with the emotional costs of this type of work and procedure.

The study also identifies the most important type of errors when correlated with its corresponding visuospatial assessments. An important error, described as the *'advancement of the needle without visualization of the needle tip'* correlated negatively with MRT. These findings are consistent with previous findings, as there exists a novice propensity to advance a needle in response to its lack of visualization on an ultrasound screen.[127,193] Furthermore, this study enumerates that better mental rotation skills are associated with decreased tendency for this type of error, which is clinically vital for completing an intervention in UGRA securely and successfully. Similarly, we found another type of error characterized as *'number of needle passes'* that correlated negatively with MRT.

Further sub-analysis of the type of performances in relation to its visuospatial assessments revealed some important findings. Participants with better mental rotation skills were related to perform better and with increased efficiency in terms of *'time and motion', 'instrument handling' and the 'flow of procedure'*.

Limitations are integral to any investigation and warrant specific comment here. Aside from the ethical problems of allowing novice practitioners to practice on real patients, it is likely that anatomical differences, doctor-patient interactions and the pressures of achieving successful blocks would generate inconsistent results in real clinical situations.[147,194] For the purpose of this study, we used a turkey-breast bench model rather than *in vivo* needling. Despite the lack of clinical context, we believe that our participants experienced an "examination-like" stress caused by their assessments during the study, and that it is likely that this stress could produce detrimental and variable effects on performance such as that in clinical practice.[169,170,172] Despite the

limitations of the turkey-breast bench model, it is accepted as an initial means to evaluate novice performance in UGRA[149] and to perform training in this complex technical task.[192] As such, we felt that our bench-top simulation provided a reproducible and realistic environment in which to assess our subjects.

The subjects were medical students and not practicing doctors, therefore one could argue that with seniority and experience there is an inherent level of psychomotor expertise, which confers an improved ability to perform new psychomotor skills. Thus, it could be debated that had we studied anesthesiology residents, the correlations between MRT and GRS or CES may have been weaker. However, previous work has demonstrated that medical student performance of an ultrasound-guided needle task is broadly comparable to that of novice resident doctors.[193,194] Hence, we do not consider that the use of medical student volunteers presents a significant limitation to our study.

We have used assessments of visuospatial ability, emotional processing and general cognitive ability, which are considered to be valid and reliable.[180-182,184-187,190] With regard to CES and GRS, we have demonstrated high levels of inter-rater agreement and internal consistency of the assessment tools; this is in line with the previous findings.[146,194] We believe that the high inter-rater agreement reflects the assessor training with the CES and GRS tools, prior to recruitment. As such, we believe that our measurement of task performance is both reliable and valid.

Lastly, we have attempted to mitigate for any bias in assessment by asking our assessors to rate the participants' performance independent of one another, and without knowledge of the participants' scores in the various psychological assessments completed beforehand.

This study has been able to pinpoint imperative features of learning and expertise development among novices. However, participants selected for the purpose of this study with high baseline visuospatial ability may restrict the sensitivity of the visuospatial assessments, which were created for a wider population, thus affecting the general results of the population.

The premise of this study was to identify whether visuospatial ability could be used to predict novice performance of an ultrasound-guided needle task. In doing so, we have identified correlations between performance, mental rotation skills, and negative mood. Future research could investigate whether specific training interventions could transform visuospatial ability and thus enhance skills acquisition in UGRA. With regard to MRT as a screening tool to focus training, we believe that future studies need to assess the sensitivity and specificity of MRT in this context.

Appendix 1

Appendix 2

The University of
Nottingham

University of Nottingham
School of Clinical Sciences
Division of Anaesthesia and Intensive Care
Queen's Medical Centre
Derby Road
Nottingham
0115 8231009

Participant Information Sheet

Visuo-spatial ability as a predictor of novice performance in ultrasound-guided regional anaesthesia.

Researchers:

Dr. A Shafqat, Prof. E Ferguson, Dr. V Thanawala, Dr. N Bedforth, Prof. J Hardman and Dr. R McCahon.

Invitation paragraph

You have been invited to take part in a research study. Before you decide whether to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends and relatives if you wish to. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you wish to take part or not. If you decide to take part you may keep this leaflet. Thank you for reading this.

Background

The ability to perform practical procedures competently and safely is an important aspect of the practice of Anaesthesia. The assessment of procedural skills in anaesthesia is generally given less importance than assessment of knowledge and judgement-based skills. Due to an international trend of decrease working hours during training, an increase focus on patient safety and service delivery and greater accountability to the public and the government, there is an increase need for a robust system to ensure competence in procedural skills in Anaesthesia.

Previous work in laparoscopic surgery, has demonstrated that visuo-spatial ability correlates positively with laparoscopic performance. In Anaesthesia, it would be advantageous to accurately predict those trainees who require increased faculty input in order to develop their expertise in ultrasound-guided regional anaesthesia.

We aim to identify which assessment of visuo-spatial ability best predicts novice performance of ultrasound-guided regional anaesthesia. This will permit early, targeted training for those individuals who are predicted to perform poorly. The study will last from August 2012 till October 2012.

Appendix 3

What does the study involve?

The study will be conducted in the University department of Anaesthesia, Queen's Medical Centre Campus, Nottingham University Hospitals NHS Trust, U.K. It is an observational study involving healthy medical students from the University of Nottingham. Potential volunteers wishing to participate will be invited to attend and informed consent will be gained on the day of the study, including consent for video recording.

The study consists of following stages below:

Stage 1: Assessment of Visuo-Spatial Ability

The volunteers' visuo-spatial ability will be assessed by a number of psychological tools. These include: Mental rotation test, Spatial reasoning (AH4) test, Numerical test, measurement of anxiety and mood and Field independence test. Assessments will be made by a psychologist trained in the use of these assessment tools in a separate room.

Stage 2: Teaching Intervention

The volunteers will be shown an 11 minute video of expert performance of ultrasound-guided needle advancement, that is mapped to specific learning objectives. The volunteers will have a total of 30 minutes to watch and review the video.

Stage 3: Ultrasound-Guided Needling Task

The volunteers will be asked to complete the needling task on a turkey breast model inserted into a manikin, using an ultrasound probe. The performance of the task will be video recorded. No feedbacks will be given to the volunteers while performing the tasks. Study participation will cease once the ultrasound task has been completed. The Volunteers will be expected to;

i) Switch on the ultrasound machine

ii) Correctly orientate the ultrasound proble (linear, 38mm) in relation to the display on the screen

iii) Ensure adequate application of conducting gel to enhance ultrasound transmission and picture quality

iv) Locate and identify the target (Olive) within the turkey breast

v) Adjust the gain function to improve the quality of the image by altering brightness of the picture

vi) Alter the depth of the image to obtain a suitable image of the target

vii) Using an in-plane approach insert a 50-mm Stimuplex A™ needle into the turkey breast and aim to place the needle tip at the 12 o' clock position, as indicated by the attending assessors, above the upper edge of the target, without piercing the target

Appendix 4

Assessment of the Ultrasound Task

Two anaesthetists, experienced in performing ultrasound-guided regional anaesthesia, will assess the volunteer's performance of the task independently. The video recording will only be used for reference if there are any uncertainties concerning the volunteer's performance.

Why have you been chosen?

You have been chosen because you are aged over 18 years, you are healthy, and you are able to give informed consent to participate in this study.

Do you have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

What are the possible disadvantages and risks of taking part?

There are no known health issues associated with the use of ultrasound. Possible side effects include accidental needle prick injury during handling. If you feel uncomfortable in any way, you can stop at any time. You will also need to be able to spare a total of 150 minutes of your total time.

What if something goes wrong?

If you should have any complaints of any aspect of the study you should initially contact;

Professor Jonathan Hardman
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Telephone 0115 8231002
E-mail J.Hardman@nottingham.ac.uk

If you remain unhappy and wish to complain formally, please contact the secretary of the University of Nottingham Medical School Research Ethics Committee, Mrs Louise Sabir, E-mail address: louise.sabir@nottingham.ac.uk

In the unlikely event that you suffer injury to yourself or damage to your property as a result in taking part in this research, the University does have an insurance policy to cover harm arising as a result of the defect in the design of the study. In addition, all medical practitioners taking part in the research have personal medical negligence cover.

Appendix 5

Will my taking part in this study be kept confidential?
All information, which is collected, about you during the course of the research will be kept on a password protected database and is strictly confidential. Any information about you that leaves the research unit will have your name and address removed so that you cannot be recognised from it. All information generated by this study will be archived securely within the University Division of Anaesthesia & Intensive Care, University of Nottingham and destroyed 7 years after the study is completed.

What will happen to the results of the research study?
The results will be presented as a written project authored by Dr Atif Shafqat in completion of his PhD. The findings of this study may also be presented at national or international meetings and may be published in a medical journal.

Who is organising and funding the research?
The study is part of post-graduate research supported by The University of Nottingham. There is no external funding.

Who has reviewed the study?
This study has been reviewed and approved by the University of Nottingham Medical School Ethics Committee.

Contact for Further Information
Dr Atif Shafqat
PhD. Student University of Nottingham
University Department of Anaesthesia
Queens Medical Centre
Nottingham NG7 2UH
Email: msxas4@nottingham.ac.uk

Many thanks for your participation in this study.

Appendix 6

The University of Nottingham

University of Nottingham
School of Clinical Sciences
Division of Anaesthesia and Intensive Care
Queen's Medical Centre
Derby Road
Nottingham
0115 8231009

PARTICIPANT CONSENT FORM

Visuo-spatial ability as a predictor of novice performance in ultrasound-guided regional anaesthesia.

Name of Researchers:
Dr. A Shafqat, Prof. E Ferguson, Dr. V Thanawala, Dr. N Bedforth, Prof. J Hardman and Dr. R McCahon.

Study Number:_____

Please write your **initials** in each box

1.   I confirm that I have read and understand the information sheet (version number **1.0** dated the **12th June 2012**) for the above study and have had the opportunity to ask questions. ☐

2.   I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that information may still be used in the project analysis. ☐

3.   I understand that all data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to me taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4.   I understand and agree that I may be video-recorded during my participation in the study and that this video recording will only be used for reference by the study investigators. ☐

5.   I agree to take part in the above study. ☐

_____   _____   _____
Name of Participant   Date   Signature

_____   _____   _____
Name of Researcher   Date   Signature

When completed: - 1 copy for participant, 1 copy for researcher site file.
Participant Consent Form. Visio-Spatial ability in UGRA. University of Nottingham. Final Version 1.0
12th June 2012.

80

Appendix 7

Visuo-spatial ability as a predictor of novice performance in ultrasound-guided regional anaesthesia (UGRA).	**Participant number**		**Study period**	

Errors - Please tick each time an error is made

Needle advanced without visualisation of needle tip	
Failure to identify the target	
Failure to recognise orientation of probe with the image screen	
Unintentional probe movement	
Target malpositioned on screen, including incorrect depth selection	
Attention focussed on hand and not image as needle advanced	

Total number of errors   **Box A**

Number of needle passes*   **Box B**

* A needle pass is defined as a new puncture of the turkey breast or if the needle is withdrawn towards the exterior of the turkey breast.

Image quality scale (select one)

Scale	Image quality	Description
0	Ideal	· Olive viewed on highest resolution & all contents visualized · Whole needle seen as contacted wall
1	Good	· Olive imaged well enough to define all aspects of contents · Needle tip & part of shaft seen as contacted wall
2	Satisfactory	· Olive visualized so at least wall was identifiable · Needle tip seen as contacted wall
3	Poor	· Olive wall was only partly visualized · Needle contact was identifiable only by tissue distortion

0	Ideal
1	Good
2	Satisfactory
3	Poor

Box C

Composite Error Score
(Box A + Box B + Box C)

Time to perform task   s

Appendix 8

Visuo-spatial ability as a predictor of novice performance in ultrasound-guided regional anaesthesia (UGRA).

Participant Number [   ]   **Study Period** [   ]

Global Rating Scale for UGRA.

Please tick and score each of the 7 items below appropriately;

	1	2	3	4	5
Preparation for procedure	Did not organise equipment well Has to stop procedure frequently to prepare equipment. [ ]	[ ]	Equipment generally organised. Occasionally has to stop and prepare items. [ ]	[ ]	All equipment neatly organised, prepared and ready for use. [ ]
Time and motion	Many unnecessary moves. [ ]	[ ]	Efficient time/motion but some unnecessary moves. [ ]	[ ]	Clear economy of movement and maximum efficiency. [ ]
Instrument handling	Repeatedly makes tentative or awkward moves with instruments. [ ]	[ ]	Competent use of instruments but occasionally appeared stiff or awkward. [ ]	[ ]	Fluid moves with instruments and no awkwardness. [ ]
Flow of Procedure	Frequently stopped procedure and seemed unsure of next move. [ ]	[ ]	Demonstrated some forward planning with reasonable progression of procedure. [ ]	[ ]	Obviously planned course of procedure with effortless flow from one move to the next. [ ]
Image quality	Inadequate image. Target partially visualised. Needle movement only identifiable by tissue distortion. [ ]	[ ]	Good image. Most of target visualised. Needle tip and part of needle shaft visible. [ ]	[ ]	Outstanding image. Unequivocal image with complete structure visualised. Whole of needle visible. [ ]
Knowledge of procedure	Deficient knowledge. [ ]	[ ]	Knew all important steps of procedure [ ]	[ ]	Demonstrated familarity with all aspects of procedure [ ]
Overall performance	Very poor. [ ]	[ ]	Competent. [ ]	[ ]	Clearly superior. [ ]

CHAPTER – 3

Construct Validity And Reliability Of An Objective Structured Assessment Tool For Performance Of Ultrasound – Guided Regional Anesthesia (UGRA)

INTRODUCTION

The introduction of the Calman system in the United Kingdom, the implementation of the European Working Time Directive, and the financial pressures to increase productivity[203] have reduced clinical opportunities during training. An increased focus on patient safety and greater accountability to the public, media and government has led to increased demands for the development of robust and objective assessment of physician's technical and psychomotor skills[204], which is required for the safe performance of invasive procedures in medicine.[143,204] The assessment of technical skills during training has been regarded as a form of quality assurance for the future.[205] Training in medicine is traditionally based on an apprenticeship model, in which the assessment of a technical expertise is a task of the senior trainers. However, their assessment is largely subjective.[206] Objective assessment is crucial because deficiencies in training and performance are hard to correct without objective feedback.[207]

The ability to perform practical procedures safely and efficiently is of paramount importance to the practice of anaesthesia. Historically though, the assessment of procedural skills in anaesthesia has generally been given less importance than the assessment of knowledge and judgement-based skills.[107,151] The last two decades have borne witness to rapid developments in the practice of ultrasound-guided regional anaesthesia (UGRA), which is a complex invasive procedure that requires both technical and non-technical expertise to perform safely, effectively and efficiently.[208,209] Specialists of this anaesthetic subspecialty along with other distinct national and international regional anaesthesia societies have published a number of

educational core curriculums and structured training guidelines to teach and clinically put into practice these new skills.[210,211] However, the development of complementary assessment tools has lagged behind. Therefore a method of evaluation is needed to ascertain the competence of junior residents in performing the new skills and accordingly the appraisers should be facilitated to provide objective feedback that will potentially assist junior residents in improving their skills.[212]

At present, technical and psychomotor expertise in anaesthesia is typically assessed by workplace-based assessments (WBA).[13] Workplace-based assessment is defined by the U.K. postgraduate medical education and training board (PMETB), which is now part of the general medical council (GMC) as:

"The assessment of working practices based on what trainees actually do in the workplace, and predominantly carried out in the workplace itself."[213]

The aim of the anaesthetic training programmes is to train doctors up to a level at which they are capable of independent specialist practice. Traditional assessments have considerably given emphasis to knowledge acquisition than clinical ability. Therefore many postgraduate and undergraduate programmes have set up workplace-based assessments. Workplace-based assessments are fundamental for any competence-based postgraduate curriculum and are now unavoidable for many specialist-training programmes.[214]

Reliability and validity are essential for any assessment, including workplace-based assessments. Despite the fact that anaesthesia fellowship examinations are valid and reliable tests of knowledge, but they may not be an appropriate assessment of the capability to practice as an anaesthetist. Workplace-based assessments are more valid measure of this aptitude, however according to few studies they have low reliability.[215,216]

In contrast with fellowship examinations, there are a number of factors, which affect the reliability of workplace-based assessments. Fellowship examinations can be homogeneous for difficulty and content, but workplace-based assessments cannot because cases are unpredictable and cannot be planned or recurring. The examiners in the anaesthesia fellowship examinations are trained, and agree on set standards of performance, on the contrary the examiners in workplace-based assessments can include all specialist anaesthetists working in teaching hospitals, many of whom have limited training in the use of the workplace-based assessment tools.

Workplace-based assessment tools such as the direct observation of procedural skills (DOPS) do not describe or quantify acceptable practice. As such, practice is deemed satisfactory, or not, based on the judgement of more experienced, senior colleagues. Whilst this approach is cheap and has face validity, it is not sufficiently robust to withstand scrutiny and would not be adequate to support credentialing in the future. The methods for assessing technical skills including logbooks and direct observation without specific criteria are based on informal global assessments of performance and are subjective and unreliable.[120,151,217] On the contrary direct observation with specific criteria, which are essential to structured learning of the expertise and may

significantly improve the assessment of a technical skill, are more objective, valid and reliable.[125,217] This includes expert-based ratings like skill-specific checklist and global rating scale (GRS), and computer-based assessments.[146,218]

A criterion-referenced checklist with global rating scale to assess the performance of ultrasound-guided regional anaesthesia has been developed using a modified Delphi method.[219] In the 1950s scientists at the Rand Corporation developed the Delphi method as a tool for making informed decisions based on expert opinion.[220] The Delphi method is a structured process for collecting and distilling knowledge from a group of experts, by means of a series of questionnaires combined with controlled opinion feedback.[221] According to this method the members or experts are encouraged to debate among each other without having the need of meeting personally. The members participate while keeping themselves anonymous. The outcome is that all opinions are given equal significance and this eliminates any influence of strong opinions from any personality. This also diminishes significantly the social-emotional behaviour that is often found when using other methods and allows the members to focus on task-oriented activities.[120] The Delphi method has been used to generate content in a variety of contexts, including the development of policies, curricula, best-evidence practice guidelines, and competency assessment tools.[222-224]

Two educators in ultrasound-guided regional anaesthesia from Sunnybrook Health Sciences Centre and the University of Toronto have formulated an initial checklist of 30 items for the performance of ultrasound-guided regional anaesthesia.[219] This pilot checklist was e-mailed to 18 reviewers for the first round of feedback in six independent academic teaching centres with dedicated regional anaesthesia teaching

programmes in Canada and the United States. All the reviewers were specialist anaesthetists and educators with more than six years of experience in this field and were involved in ultrasound-guided regional anaesthesia from its early origin years. The experts were requested to score each item on the checklist as ''yes'' or ''no'' for inclusion or exclusion in the list respectively. They were also asked to provide any additional items, which were relevant and further considered necessary to be included in the checklist. In addition to that, the experts were encouraged to rate the importance of each item on the checklist on a five-point Likert scale (0 = extremely irrelevant, 1 = irrelevant, 2 = neutral, 3 = important, 4 = extremely important). They were also given the opportunity to provide comments on the checklist. The results were then anonymously collated, and the checklist was modified on the basis of the feedback received from the reviewers. Later it was sent back to the same reviewers for second round of feedback.

After the second round of feedback, a number of reviewers recommended that items relevant to general patient care during ultrasound-guided regional anaesthesia should be removed from the checklist and be added into a modified global rating scale (GRS). Thus items, which were relevant to technical aspects of ultrasound-guided regional anaesthesia, were included into the checklist. In order to assess the performance in the nontechnical aspects of ultrasound-guided regional anaesthesia a global rating scale was created. The global rating scale was adapted from the original objective structured assessment of technical skills (OSATS).[120] The global rating scale, which was created contains nine categories scored on a five-point Likert scale, with descriptions for the extreme and median scores specific to regional anaesthesia. Seven of these categories were adopted from the original objective structured

assessment of technical skills (OSATS)[120], and two new ones were added, namely, ''asepsis'' and ''patient interaction,'' in order to compensate for the removal of the relevant items from the checklist.

The checklist and global rating scale was then re-distributed to the reviewers for further consideration. No further modification was recommended in the third round of feedback and thus the consensus was developed. This number of iterations is consistent with other medical studies that have applied the Delphi method.[225] Ultimately, the assessment tool consisting of a final assessment checklist (AC) containing 22 items and global rating scale (GRS) with nine categories for the performance of ultrasound-guided regional anaesthesia was formulated by the Delphi method.[219]

This study[219] established high face validity and concurrent validity of the assessment tool. However, it did not attempt to demonstrate the construct validity or reliability of the assessment tool. Therefore, we aim to examine the construct validity and reliability of both components of this new assessment tool for the performance of ultrasound-guided regional anaesthesia, by examining whether it can adequately differentiate between performance levels in anaesthetists across the spectrum of expertise.

METHODS

The study was reviewed and approved by the University of Nottingham Medical School Research Ethics Committee (Approval Reference; K09052013LT 13053 SCS Anesthesia).

Design:

This was a single centred, qualitative, observational study, which used a convenience sample across one NHS Trust in England. The study was conducted at the University Department of Anaesthesia, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham UK. The aim was to investigate the validity and reliability of a criterion-referenced assessment tool for the performance of ultrasound-guided regional anaesthesia that will determine whether or not the assessment tool could discriminate between different levels of expertise in anaesthetists performing ultrasound-guided regional anaesthesia.

Subjects:

The subjects were all grades of anaesthetists who performed ultrasound-guided regional anaesthesia as their usual clinical practice, which included all patients who were having ultrasound-guided regional anaesthesia as part of their planned anaesthetic management. A participant information sheet was forwarded to those anaesthetists who expressed an interest via e-mail or in person. Potential anaesthetists wishing to participate were invited to attend and informed consent was gained in advance of any clinical activity. The patients of the anaesthetists who had consented

to take part in the study were consented on the morning of their day of surgery. Two independent expert anaesthetists experienced in performing ultrasound-guided regional anaesthesia objectively assessed any single performance of ultrasound-guided regional anaesthesia by all the anaesthetists on their patients.

Setting:

Anaesthetists performing ultrasound-guided regional anaesthesia on their patients were observed in the following locations:

1. The Block room, Main Operating Theatre Suite, Queen's Medical Centre, Nottingham University Hospitals NHS Trust.

2. Anaesthetic rooms in theatre suites, Nottingham University Hospitals NHS Trust.

Recruitment:

A poster [appendix 9] advertising the study was distributed by the departmental secretarial staff through e-mail to all the Anaesthetists including consultants, non-career grades and trainees who were working at Nottingham University Hospitals NHS Trust. A copy of the poster was also displayed in the following areas:

- Block room, Main Theatres, Queen's Medical Centre

- Coffee room, Department of Anaesthesia, Queen's Medical Centre

- Common room, Department of Anaesthesia, Nottingham City Hospital

- Orthopaedic & Trauma Theatres, Queen's Medical Centre

- Orthopaedic Theatres, Nottingham City Hospital

Two information leaflets were created:

➢ *One for anaesthetists [appendix 10-13];* this was distributed via the Trust internal mail and emailed to all anaesthetists informed working at Nottingham University Hospitals NHS Trust. Consent [appendix 14] was then gained from Anaesthetists who wished to take part in advance within 7 days before any clinical activity.

➢ *One for patients [appendix 15-18];* this was given to all the patients of anaesthetists who had consented to take part in the study. This information leaflet was given to patients on the morning of their surgery. They were already consented for their surgery and their anaesthetic management; so therefore consent [appendix 19] for participating in the study was gained within 1 – 4 hours before their surgery by the participating anaesthetist and / or the research investigators.

Inclusion criteria:

Anaesthetists:

1. Any anaesthetist who planned to perform ultrasound-guided regional anaesthesia from whom informed consent has been obtained.
2. Ability to give written informed consent.
3. Aged 18 years or older (no upper age limit).

Patients:

1. Undergoing ultrasound-guided regional anaesthesia.
2. Ability to give written informed consent.
3. Aged 18 years or older (no upper age limit).

1. Anaesthetists and patients who did not undergo ultrasound-guided regional anaesthesia. The clinical decision to perform ultrasound-guided regional anaesthesia was taken by the attending Anaesthetist.

2. Lack of informed consent from either the attending Anaesthetist or the patient precluded entry into the study.

Study Regimen:

Written informed consent was obtained from the recruited anaesthetists and their patients as described above. The study took place during the routine operating lists and the anaesthetist performed ultrasound-guided regional anaesthesia, as they normally did so on their patients. The researchers observed the actions of the anaesthetist while they were using ultrasound to perform regional anaesthesia on their patients. They observed single performance of any ultrasound-guided regional anaesthesia by the participant anaesthetist, which formed part of the planned anaesthetic management of their patients.

The two expert examiners objectively assessed all the performances of ultrasound-guided regional anaesthesia. This assessment was in real time during the performance of the ultrasound-guided regional anaesthesia and began with the initial preparation and equipment setup and ended at completion of the procedure. The experts did not influence the clinical practice of the participating anaesthetist in any way. For each individual case, the observations stopped when the patient left the anaesthetic room. Each observation took the length of the ultrasound-guided

regional anaesthesia, which was about 20 minutes on average. Patients' participation was for the duration of the ultrasound-guided regional anaesthetic procedure.

Following thorough expert consensus by the research team, the participating anaesthetists were asked compulsory demographic information [appendix 20] regarding their level of training or seniority in terms of their ultrasound-guided regional anaesthetic practice. For the purpose of analysis and based on the compulsory demographic information regarding their ultrasound-guided regional anaesthetic practice in the preceding year, the participants were divided into 3 groups.

Group – I: consisted of anaesthetists who performed *30 or less* ultrasound-guided nerve blocks in the preceding year.

Group – II: comprised of anaesthetists who did between *31 – 100* ultrasound-guided nerve blocks during the previous year of their practice.

Group – III: entailed anaesthetists who completed *more than 100* ultrasound-guided nerve blocks in the past year of their practice.

The participant anaesthetists were also asked to self-rate themselves for their ultrasound-guided regional anaesthesia expertise as; *novice, competent or expert*.

Blinding:

Participants were given an identification number following recruitment, which was recorded in the participant identification log. Both the expert examiners were blinded to the participant's seniority or level of training in terms of their ultrasound-guided regional anaesthetic practice.

1. One of the experts had not worked clinically at Nottingham University Hospitals NHS Trust. As such, he was not familiar with any of the consultant, career grade or trainee anaesthetic staff.

2. The other expert, although worked in the same NHS Trust but was not familiar with the ultrasound-guided regional anaesthetic practice of the participating anaesthetists.

Therefore to minimise the observer bias, both the expert examiners were kept blinded about the compulsory demographic information regarding the ultrasound-guided regional anaesthetic practices of the participating anaesthetists. This information was only disclosed for analysis after completion of all the performances at the end of the study.

Assessment tool:

The assessment tool[219] comprised of two components:

1. An assessment checklist (AC) for ultrasound-guided regional anaesthesia.

2. A global rating scale (GRS) for ultrasound-guided regional anaesthesia.

The assessment checklist [appendix 21] for ultrasound-guided regional anaesthesia has 22 items which are assessed using a 3-point scale (not performed = 0, poorly performed = 1, well performed = 2).

The global rating scale [appendix 22] for ultrasound-guided regional anaesthesia consists of nine categories, scored on a 5-point Likert-scale each of which has performance descriptors to aid assessment. The expert examiners had undergone specific training and practice in the use of these assessment tools.

Statistics:

Demographic information:

> ➢ Ultrasound-guided regional anaesthetic practice:

   1) Total number of ultrasound-guided nerve blocks performed in the preceding year.

   2) Total number of ultrasound-guided nerve blocks performed in the preceding month.

   3) A self-assessment of the ultrasound-guided regional anaesthesia expertise as; novice, competent and expert.

Primary outcome:

Total (Mean) scores of the assessment checklist and global rating scale.

Secondary outcome:

1. Inter-rater agreement for the assessment checklists and global rating scale.

2. Internal consistency of the assessment checklists and global rating scale.

Before starting the analysis, we checked the data for errors and missing values and gained an overview. Descriptive statistics for demographic and outcome measure data was calculated.

For continuous where appropriate, outcome data were transformed to conform to normality. Where normality could not be established, non-parametric analysis was conducted. The data were assessed for goodness-of-fit to the normal distribution using the Kolmogorov–Smirnov test. All statistical analysis was performed using STATA/IC version 10.0 (StataCorp, Texas, USA). Univariate description was presented for all the variables. One-way analysis of variance (ANOVA) test was used to compare the means of normally distributed outcome data from the 3 groups.

Assessment Tool Reliability:

Reliability was assessed by the following tests:

1) The inter-rater agreement for the assessment checklist and global rating scores was calculated by the Pearson correlation coefficient.

2) The internal consistency (the inter-item agreement) of the assessment checklist and the global rating scale was assessed by Cronbach's alpha.

Sample size and justification:

The study was conducted at Nottingham University Hospitals NHS Trust. A single performance of all the anaesthetists participating in the study was observed during the study. In line with previous studies[106,117,146] we aimed to recruit between 20 to 40 participant anaesthetists.

RESULTS

Demographics:

21 anesthetists were enrolled in the study. Demographic information regarding UGRA practice included type of blocks performed, total number of UGRA blocks performed in the previous year and month respectively and self-assessment of the UGRA expertise.

Different types of UGRA blocks were performed and assessed during the study. The two most common type of UGRA block performed by the participant anaesthetists were femoral (42.86 %) and interscalene (33.33 %), respectively (table 6). This was followed by axillary (9.52 %). Other UGRA blocks included ankle, popliteal and rectus sheath catheter (table 6).

Table 6 Types of UGRA blocks performed

| Types of UGRA block	N (%)
Ankle	1 (4.76)
Axillary	2 (9.52)
Femoral	9 (42.86)
Interscalene	7 (33.33)
Popliteal fossa	1 (4.76)
Rectus sheath catheter	1 (4.76)

UGRA – Ultrasound-guided regional anaesthesia

The total number of UGRA blocks performed by the participating anaesthetists in the '*preceding year*' were minimum = 5, maximum = 600, median = 90 and interquartile range (IQR) = 130.

Similarly, the total number of UGRA blocks completed by the participating anaesthetists in the '*past month*' were minimum = 1, maximum = 45, median = 10 and interquartile range (IQR) = 15.

Out of the total participant anaesthetists, 61.9 % rated themselves as competent, whereas 19.05 % each rated themselves as novices and expert, respectively (figure 8).
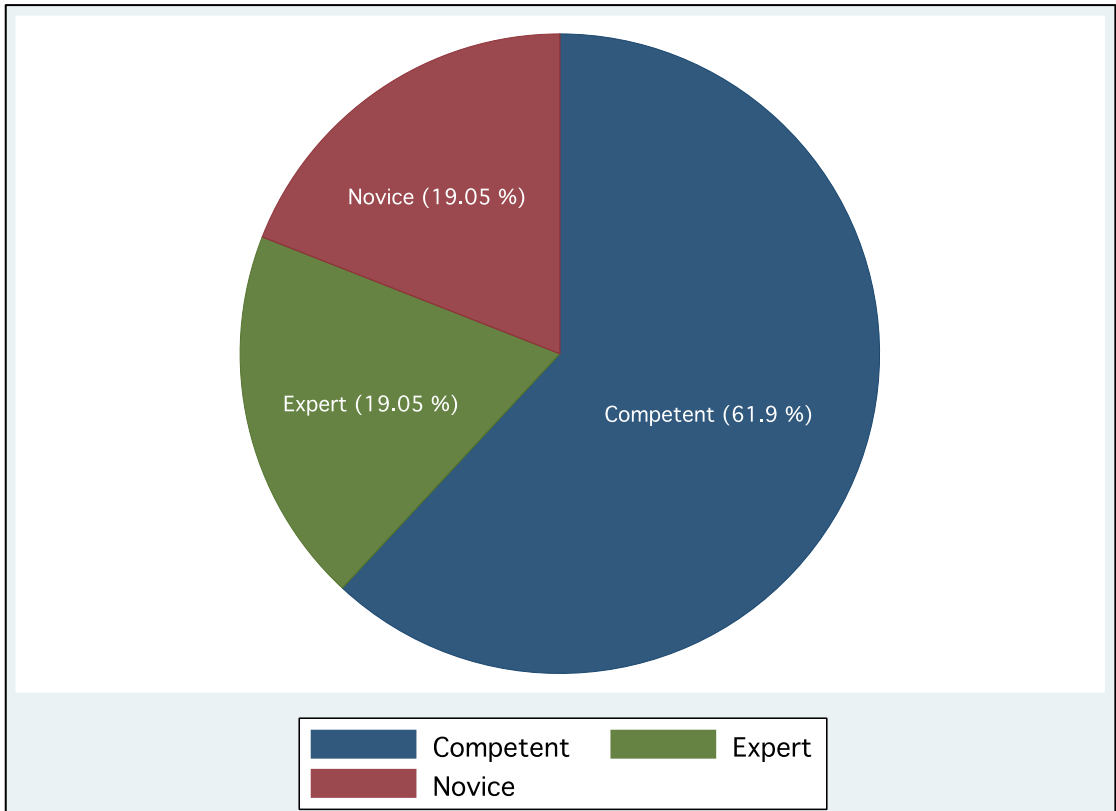
Figure 8 Self-Assessment of UGRA expertise by all study participants

Further sub-analysis revealed self-assessment of the UGRA expertise by all anaesthetists according to each of the three groups (table 7).

Table 7 Self-Assessment of UGRA expertise according to the three study groups

|	Novice N (%)	Competent N (%)	Expert N (%)
Group – I	2 (28.57)	5 (71.43)	0 (0.00)
Group – II	1 (16.67)	5 (83.33)	0 (0.00)
Group – III	1 (12.50)	3 (37.50)	4 (50.00)

Group I – Anaesthetists performing ≤ 30 UGRA nerve blocks in the preceding year

Group II – Anaesthetists performing 31-100 UGRA nerve blocks in the preceding year

Group III – Anaesthetists performing > 100 UGRA nerve blocks in the preceding year

UGRA – Ultrasound-guided regional anaesthesia

It may be noted (table 7) that in 'Group – I' 71.43 % of the anaesthetists self-rated themselves as competent, followed by 28.57 % of those who self-rated as novices. In 'Group – II' 83.33 % of the participating anaesthetists self-rated as competent, with 16.67 % of those who self-rated as novices, respectively. None of the anaesthetists in both 'Group – I' and 'Group – II' respectively, self-rated themselves as expert. In 'Group – III' 50 % of the anaesthetists self-rated as expert, trailed by 37.5 % of those who self-rated as competent. The remaining 12.5 % of the participating anaesthetists in 'Group – III' self-rated themselves as novices.

Assessment checklist (AC):

Scores for the AC were mean (SD) = 32.85 (5.22), minimum = 22 and maximum = 40. One-way analysis of variance (ANOVA) revealed an overall significant difference in the mean AC scores among the three groups (F = 12.01, P = 0.0005). We did not measure the differences occurred between the groups. 'Group – III' had the highest AC score with mean (SD) = 37.37 (1.76) versus 'Group – II' mean (SD) = 32.16 (2.31) and 'Group – I' having the lowest mean (SD) = 28.28 (5.55), respectively (figure 9). This proved that higher total AC scores were associated with greater number of UGRA blocks performed (figure 10).
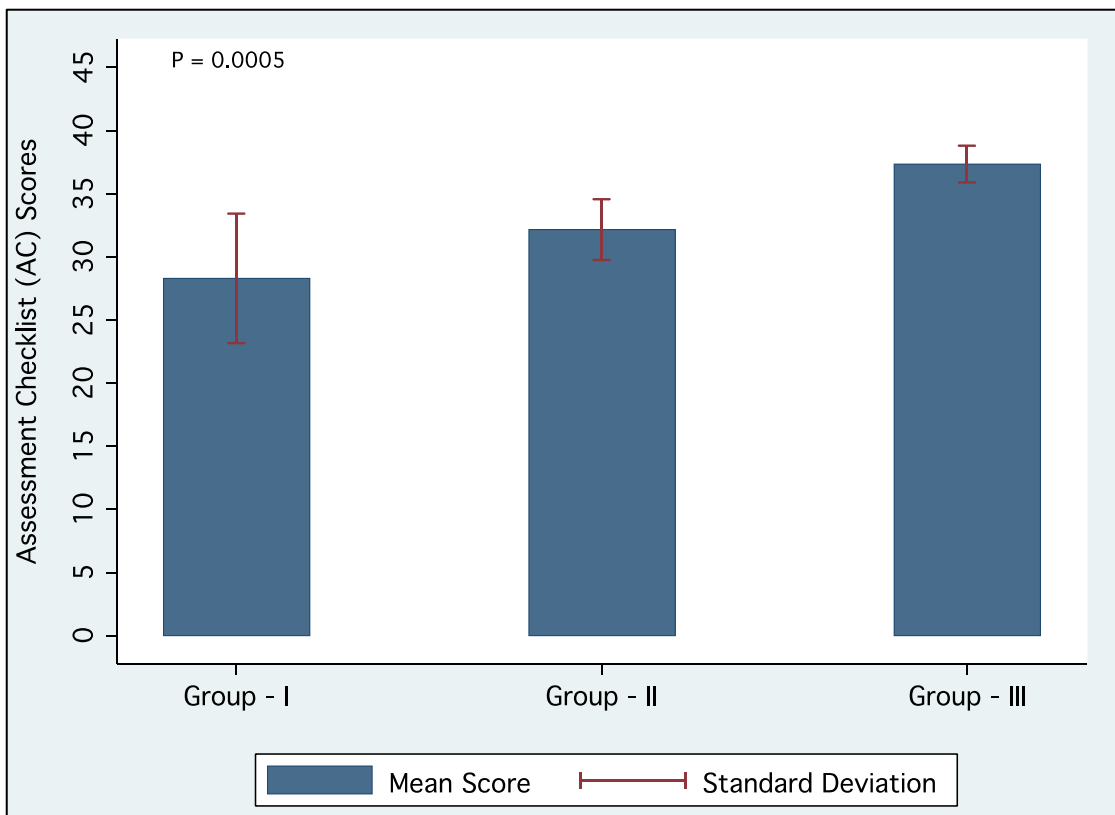
Figure 9 Assessment Checklist (AC) scores of the three study groups
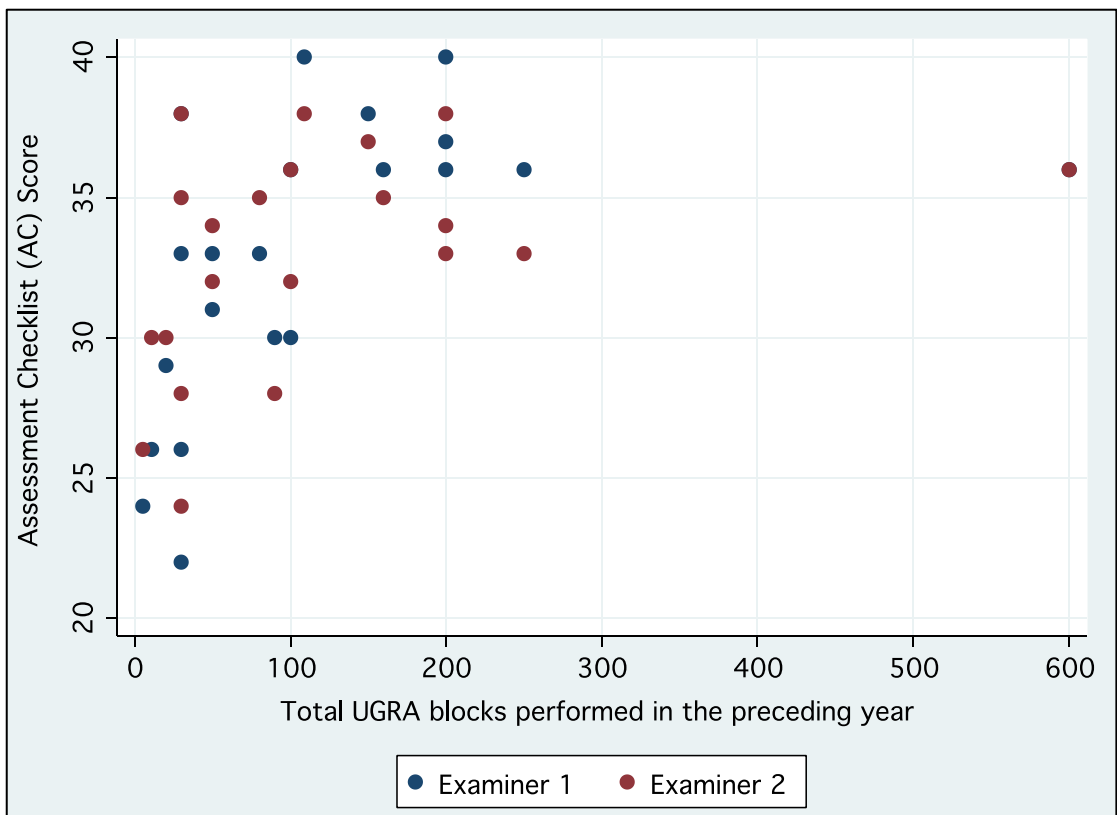
Figure 10 Assessment Checklist (AC) score *Vs*. UGRA blocks performed in the preceding year

Global rating scale (GRS):

Scores for the GRS were mean (SD) = 33.71 (8.19), minimum = 17 and maximum = 45. One-way analysis of variance (ANOVA) revealed an overall significant difference in the mean GRS scores among the three groups (F = 7.81, P = 0.0036). We did not measure the differences occurred between the groups. 'Group – III' had the highest GRS score with mean (SD) = 40.12 (5.22) versus 'Group – II' mean (SD) = 32.66 (7.86) and 'Group – I' having the lowest mean (SD) = 27.28 (6.04), respectively (figure 11). This further confirmed that higher total GRS scores were associated with higher number of UGRA blocks performed (figure 12).
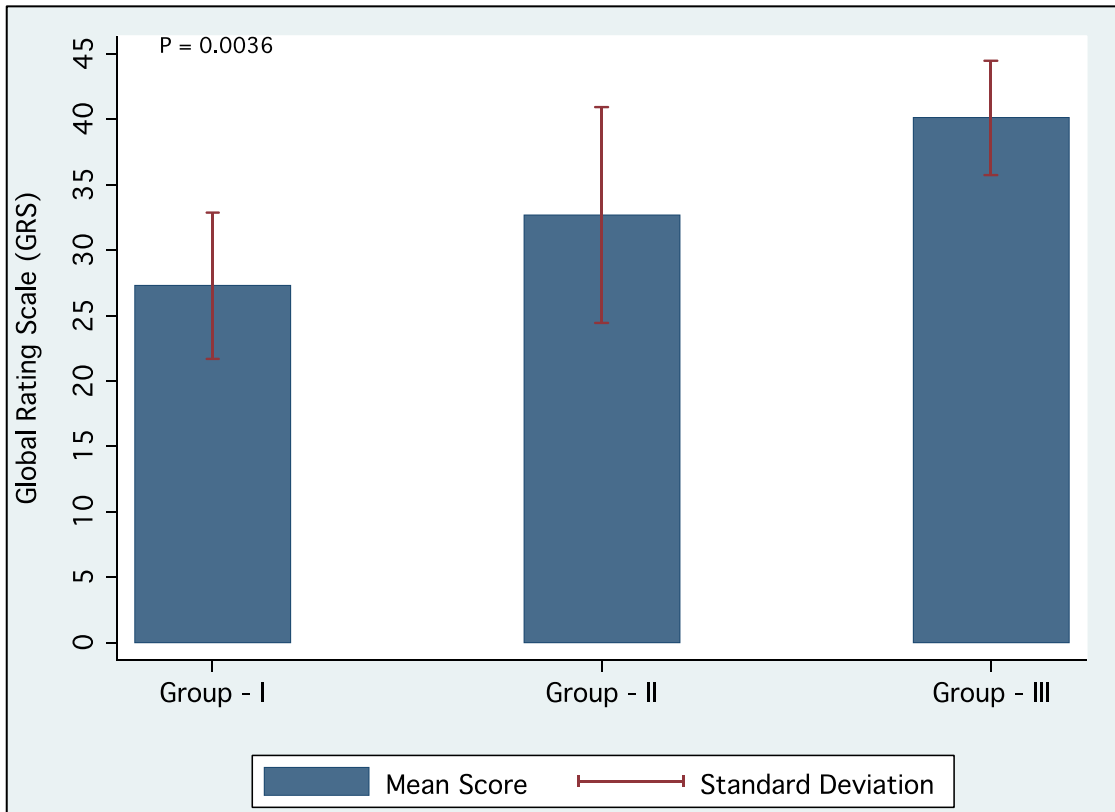
Figure 11 Global Rating Scale (GRS) of the three study groups
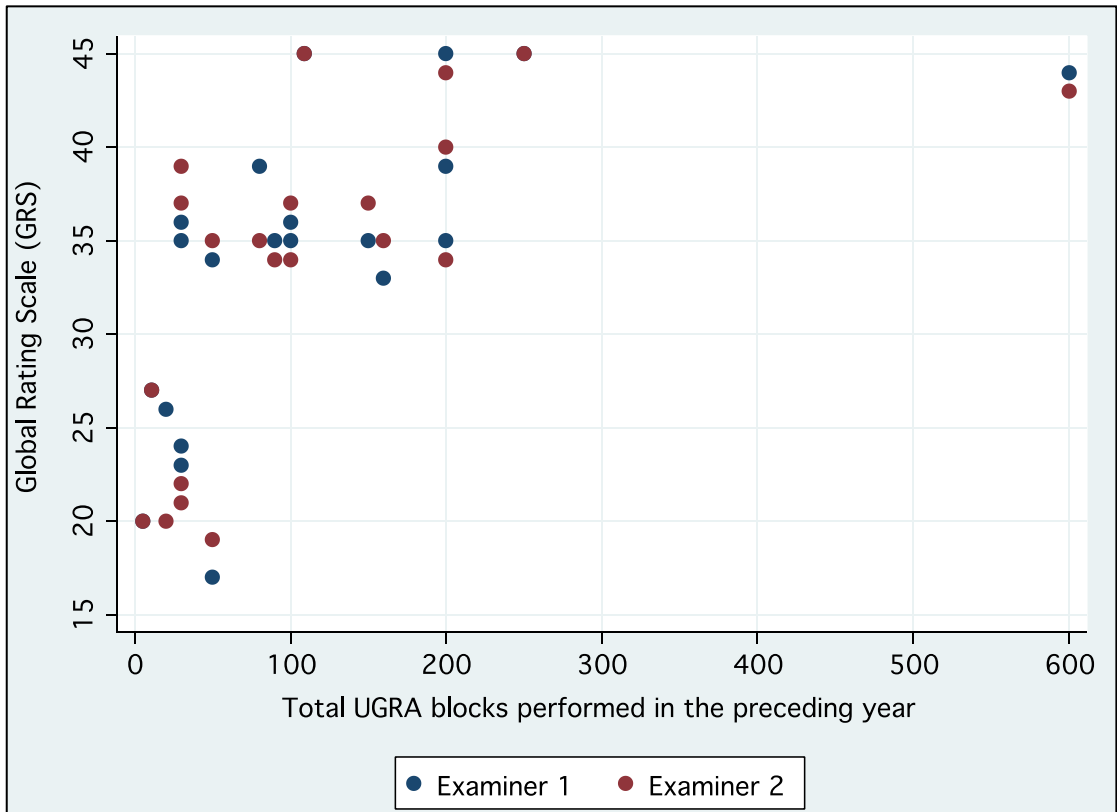
Figure 12 Global Rating Scale (GRS) *Vs.* UGRA blocks performed in the preceding year

Reliability of the assessment tool:

There was excellent inter-rater agreement for both the assessment checklist (AC) and the global rating scale (GRS). The Pearson's correlation coefficients (r) were 0.93 ($P < 0.001$) and 0.96 ($P < 0.001$) for the AC and GRS, respectively.

Similarly, Cronbach's alpha coefficient for the AC and GRS was 0.94 and 0.83 respectively, which demonstrates high inter-item consistency.

DISCUSSION:

The primary aim of this study was to establish the construct validity of a criterion-referenced assessment tool in ultrasound-guided regional anaesthesia. The results of our study illustrate a significant relationship between the number of ultrasound-guided regional anaesthetic blocks performed and the scores attained in the 2 components of the assessment tools, consisting of the assessment checklist and global rating scale respectively. Thus it concludes that the assessment tool is able to discriminate between different levels of experience and expertise of anaesthetists in the performance of ultrasound-guided regional anaesthesia. The assessment tool also reveals almost perfect agreement between the two expert examiners. The excellent inter-rater agreement and high inter-item consistency further explains that the reliability of the assessment tool is very strong.

Data on acquisition of technical skill proficiency in anaesthesia are scarce.[226] Even though, objective and reproducible systems have become the norm for expertise appraisal but their application in anaesthesia has always lagged behind. Regional anaesthetic procedures are complex to learn as compared with the basic manual skills necessary for general anesthesia[117] and this in line with previous studies, which have confirmed that regional anaesthetic procedures are the most difficult to master.[226,227]

In the United States one of the prerequisites to board certification in addition to the requirement of written and oral examinations is a minimum number of procedures to be performed by the potential applicants.[151] In anaesthesia, previous studies have

endeavoured to enumerate a minimum number of procedures, whether successful or not, which are required to achieve consistency in regional anaesthesia.[134,226-228] But this attainment of a proposed number of procedures requires validity and may not guarantee proficiency since trainees might recurrently perform procedures erroneously.[151] This system only records experience to procedures and not competence or success rate. Moreover, this system is dependent on expert consensus instead of a valid assessment system or experimentation to ascertain the acceptable number of procedures that should be achieved to obtain privileges.[229]

Statistical analysis has been employed in the past to establish competence in performance of labour epidural analgesia[134] and other anaesthetic procedures.[132,230] However, this method only takes success rate into account for competency. It does not take into account all the other essential characteristics of the performance including, sterility, safety and dexterity that are significant to achieve competency.

Essential objective parameters of satisfactory performance consistently improve appraisal of technical skills. These measures make sure that trainees are being assessed on their aptitude to execute a task rather than the existing rapport between the candidate and preceptor.[125] The checklists change examiners into observers of an expertise, in contrast to interpreters of a performance, in this way remove subjectivity. The global rating scale guarantees that the precision calculated by the checklists is not decoded as competence and give assessors the chance to give an impression of a performance regardless of whether the checklists criteria were met.[125] Both instruments together provide the most valid and reliable assessment of an expertise.

In the past distinct assessment tools have been developed for a variety of regional anaesthesia techniques but all of these tools were task specific and intended for exclusive type of regional anaesthetic nerve blocks.[106,117,146] On the contrary the assessment tool[219], which was created by the Delphi method, and now being validated is the one that contains more comprehensive content in order to appraise more generalise competencies of any type of ultrasound-guided regional anaesthetic block. It is not task specific and for that reason anticipated to be used for different types of ultrasound-guided regional anaesthetic techniques.

One limitation of this study was that some of the included items of the assessment tool might not essentially be evidence based since it was created with the Delphi's methodology and where the consensus was reached from different reviewers' opinions. Another disadvantage of this assessment tool is that the process is labor intensive and demands the commitment of expert assessors.

One of the weaknesses of a 'yes / no' checklist is its propensity to be quantitative only. The authors of the study[219] have tried to include the qualitative aspect in the checklist by adding a 3-scale feature that facilitated the assessors to appraise the procedure performances more precisely.

Another weakness of the checklist is that all the items are given equivalent significance, which may result in circumstances where a high score is achieved, even if essential items are ignored.[117] Adding a 'pass / fail' option to the scoring system could compensate this problem by permitting the assessors to assign a 'fail' grade to a

skill performance if they think that an essential item or step of the procedure is not satisfactorily performed or missed.

The global rating scale may be slightly more subjective than the checklist because it reveals what the assessors believe is the acceptable flow of the procedure, in contrast to the more objective nature of the checklist.

One may argue that the assessment tool is intended to assess the process of ultrasound-guided regional anaesthesia not its outcome. While clinical outcomes are essential conditions for assessing competence, they are not inevitably sensitive or precise measures of technical aptitude.[146] Here the issue appears that whether enhanced technical aptitude and improved course of action generate better outcomes in the form of improved clinical accomplishment and less complications. Instincts recommend this should be true but further evaluation via future studies is a requisite, particularly if formal competency assessment and credentialing is being considered.

Training programs in medicine are becoming liable to different regulatory bodies to make sure that physicians in training are objectively assessed for their skill acquirement. The validation of a criterion-referenced assessment tool[219] is a next step towards achieving this goal. Besides this, the assessment tool could also be utilized in diagnosing the deficiencies in any training program and in particular of any individual trainee who might need further teaching and training. Similarly, this may improve the way the trainees are being appraised and recognize phases of training that needed further improvement. Later the same tool can be use to record trainee's progress during the training.

The assessment of training in ultrasound-guided regional anaesthesia is moving away from the traditional informal assessment of technical skill to more specific and objective evaluations of the specific tasks that define competency[219], as demonstrated by the guidelines of American (ASRA) and European (ESRA) societies of regional anesthesia.[231] The objective assessment tools for technical skills such as assessment checklists and global rating scale, will eliminate bias and subjectivity by turning examiners into observers rather than interpreters of behaviour.[217] The assessment tool, which was created by Cheung, J.J. et al.[219] can also be utilized to identify gaps in learning and provide objective feedback that will facilitate trainees to improve their skill acquirement. The most important aspect of this assessment tool is that it has the potential to be generalised for all types of ultrasound-guided regional anaesthesia procedures.

Future study is required to evaluate the efficacy of the assessment checklist and global rating scale as training tools in anaesthesia residency programmes. In addition to that these assessment tools when used in different programs should be further studied for its learning outcomes and it's feasibility for use in the form of feedbacks from students and instructors should be examined.

Appendix 10

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Nottingham University Hospitals **NHS**
NHS Trust

School of Clinical Sciences
Division of Anaesthesia and Intensive Care
University Hospital
Queen's Medical Centre
Nottingham
NG7 2UH
0115 8231009

Participant Information Sheet - Anaesthetist
(Final Version 1.0. Date 17th April 2013)

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Name of Researchers: Dr Atif Shafqat, Dr Vishal Thanawala, Dr Kambasi Mohammed Rafi, Dr Nigel Bedforth, Dr Rob McCahon and Professor Jonathan Hardman.

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions you have. Talk to others about the study if you wish. Ask us if there is anything that is not clear.

What is the purpose of the study?

The international trend of decreasing working hours during training, an increase in focus on patient safety and greater accountability to the public and government has led to increase demands for objective assessment of manual skills and competency in the performance of invasive procedures in medicine.

The ability to perform practical procedures efficiently and safely is an important aspect of the practice of anaesthesia. The assessment of procedural skill in anaesthesia is generally given less importance than assessment of knowledge and judgement-based skills. The use of ultrasound by anaesthetists has increased dramatically. The last 2 decades have seen a rapid development in ultrasound-guided regional anaesthesia and is partly responsible for the resurgence of regional anaesthesia. Ultrasound-guided regional anaesthesia is relatively a complex invasive procedure. It comprises of a multitude of crucial technical and non-technical components, which are required in order to perform safely, effectively and efficiently. Therefore, a clear need exists to develop a reliable assessment of the performance of ultrasound-guided regional anaesthesia.

At present, the methods used to assess technical skills in Anaesthesia, such as log books and direct observation of procedural skills (DOPS) without specific assessment criteria are subjective and unreliable, since they are based on the judgement of more experienced and senior colleague. Direct observation with criteria, specifically checklists and global rating scales, may greatly improve the assessment of a technical skill. They are crucial to the structured learning of these skills. These assessment tools have previously been validated for certain types of regional anaesthesia techniques, intended to assess competencies related to a specific task.

Appendix 11

A criterion-referenced checklist with global rating scale to assess the performance of different types of ultrasound-guided regional anaesthesia techniques has been developed using a modified Delphi method. This assessment tool has already established high face validity and concurrent validity. However, it did not attempt to demonstrate the construct validity or reliability of the assessment tool. Therefore, we aim to examine the construct validity and reliability of both components of this new assessment tool for the performance of ultrasound-guided regional anaesthesia, by examining whether it can adequately differentiate between performance levels in Anaesthetists across the spectrum of expertise.

The purpose of this study is primarily educational and it will form part of the PhD. degree. In addition to that it will also determine the construct validity and reliability of a criterion-referenced assessment tool for the performance of ultrasound guided regional anaesthesia. This is an observational study, which will be conducted at Nottingham University Hospitals NHS Trust. The study will involve all grades of anaesthetists who are performing ultrasound-guided regional anaesthesia as part of their usual clinical practice. This will necessarily include any patient who is having ultrasound-guided regional anaesthesia as part of their planned anaesthetic management.

Why have I been invited?

You are being invited to participate in this study because you are a Consultant Anaesthetist or a Trainee, working within the Specialist Support Directorate of Nottingham University Hospital's NHS Trust; whose clinical decision is to perform ultrasound-guided regional anaesthesia on your patient. We are inviting approximately 50 participants like you to take part

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. This would not affect your legal rights.

What will happen to me if I take part?

Poster advertising the study will be distributed by e-mail to all anaesthetists like you, working at Nottingham University Hospitals NHS Trust.

Potential anaesthetists like you, wishing to participate will be invited to attend and informed consent will be gained in advance of any clinical activity. The patients of the participating anaesthetists who have consented to take part in the study, although are not direct participants of the study but will be consented on the day of the surgery.

The study will take place during the routine operating lists. The participating anaesthetist will perform ultrasound-guided regional anaesthesia, as they would normally do so on their patients. A single performance of ultrasound-guided regional anaesthesia from all the anaesthetists participating in the study will be observed.

Two experts will objectively assess all the performances of ultrasound-guided regional anaesthesia independently and will not influence the clinical practice of the participating anaesthetist in any way. This assessment will be in real time during the performance of the ultrasound-guided regional anaesthesia and will begin with the initial preparation and equipment setup and end at completion of the procedure.

Appendix 12

Expenses and payments

Participants will not be paid to participate in the study.

What are the possible disadvantages and risks of taking part?

There should be no disadvantages of taking part in this study. You will be observed as part of the study, if you feel uncomfortable in any way, you can stop at any time.

What are the possible benefits of taking part?

We cannot promise the study will help you but the information we get from this study may help to determine the validity and reliability of the assessment tool. After validation, the main benefit of this tool to the participants and other future anaesthetists will be that it can be used to detect gaps in learning and provide objective feedback that will enable anaesthetists to improve their skills acquisition.

What happens when the research study stops?

The study will only be stopped prematurely if we fail to recruit participants.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers contact details are given at the end of this information sheet. If you remain unhappy and wish to complain formally, you can do this by contacting patient advice and liaison service (PALS) at Nottingham university hospitals NHS trust. The contact details are given at the end of this information sheet.

Will my taking part in the study be kept confidential?

We will follow ethical and legal practice and all information about you will be handled in confidence.

If you join the study, some parts of the data collected for the study will be looked at by authorised persons from the University of Nottingham who are organising the research. They may also be looked at by authorised people to check that the study is being carried out correctly. All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.

All information, which is collected, about you during the course of the research will be kept **strictly confidential**, stored in a secure and locked office, and on a password protected database. Any information about you that leaves the hospital will have your name and address removed (anonymised) and a unique code will be used so that you cannot be recognised from it.

All research data will be kept securely for 7 years. After this time your data will be disposed of securely. During this time all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team will have access to your personal data.

Appendix 13

What will happen if I don't want to carry on with the study?

Your participation is voluntary and you are free to withdraw at any time, without giving any reason, and without your legal rights being affected. If you withdraw then the information collected so far cannot be erased and this information may still be used in the project analysis.

What will happen to the results of the research study?

The results will be presented as a written project authored by Dr Atif Shafqat in completion of his PhD. The findings of this study may also be presented at national or international meetings and may be published in a medical journal.

Who is organising and funding the research?

This research is organised by the University of Nottingham and is being funded by the institutional funds. There is no external funding for this study.

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by **Berkshire Research Ethics Committee**.

Further information and contact details

Professor Jonathan Hardman
Chief Investigator
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Tel: 0115 823 1002
Email: J.Hardman@nottingham.ac.uk

Dr Atif Shafqat
PhD. Student Researcher, University of Nottingham
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Tel: 0115 823 1004
Email: msxas4@nottingham.ac.uk

Patient Advice and Liaison Service (PALS)
NUH NHS Trust, c/o PALS,
Freepost, NEA 14614,
Nottingham NG7 1BR
Email: pals@nuh.nhs.uk
Fax: 0115 875 4655
Tel: 0800 052 1195 (City Hospital campus)
     0800 183 0204 (QMC campus)

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Participant Information Sheet. Final Version 1.0. Date 17th April 2013

Appendix 14

The University of Nottingham

Nottingham University Hospitals NHS
NHS Trust

CONSENT FORM – Participant Anaesthetist
Final version 1.0. Date 17th April 2013

Study Ref: _____

Name of Participant Anaesthetist:

Title of Project:
Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Name of Researchers:

Dr. A Shafqat, Dr. V Thanawala, Dr. KM Rafi, Dr. N Bedforth, Dr. R McCahon and Prof. J Hardman.

Please initial all boxes

1. I confirm that I have read and understand the information sheet (Final Version number **1.0** dated the **17th April 2013**) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that information may still be used in the project analysis.

3. I understand that the data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential.

4. I agree to take part in the above study.

_____   _____   _____
Name of Participant   **Date**   **Signature**

_____   _____   _____
Name of Researcher   **Date**   **Signature**
taking consent.

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Participant Anaesthetist Consent Form. Final Version 1.0. Date 17th April 2013
When completed: - 1 copy for Participant Anaesthetist, 1 (original) copy for researcher site file.

Page 1 of 1

Appendix 15

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Nottingham University Hospitals **NHS**
NHS Trust

School of Clinical Sciences
Division of Anaesthesia and Intensive Care
University Hospital
Queen's Medical Centre
Nottingham
NG7 2UH
0115 8231009

Patient Information Sheet
(Final Version 1.0. Date 17th April 2013)

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Name of Researchers: Dr Atif Shafqat, Dr Vishal Thanawala, Dr Kambasi Mohammed Rafi, Dr Nigel Bedforth, Dr Rob McCahon and Professor Jonathan Hardman.

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions you have. Talk to others about the study if you wish. Ask us if there is anything that is not clear.

What is the purpose of the study?

Like any other clinical specialty where patient safety is given prime importance through constant checks and assessments, anaesthesia (loss of sensation) requires assessment of both technical and non-technical skills that are performed on patients. Ultrasound-guided regional anaesthesia is a complex invasive procedure that requires skills to perform safely, effectively and efficiently. Therefore, there is a need to revise and test the already available assessment tool for the performance of ultrasound-guided regional anaesthesia.

The purpose of this study is primarily educational and it will form part of the PhD. degree. In addition to that it will also determine the validity and reliability of an assessment tool for the performance of ultrasound-guided regional anaesthesia. The objective is to establish that this tool can reliably assess performance of ultrasound-guided regional anaesthesia and thus differentiate between different levels of expertise.

This is an observational study, which will be conducted at Nottingham University Hospitals NHS Trust. The study will involve all grades of anaesthetists who are performing ultrasound-guided regional anaesthesia as part of their usual clinical practice. This will necessarily include you (any patient) who is having ultrasound-guided regional anaesthesia as part of your planned anaesthetic management. Although you are not direct participant, but will be consented on the day of your surgery.

The participating anaesthetist will perform ultrasound-guided regional anaesthesia, as they would normally do so on their patients. A single performance of ultrasound-guided regional anaesthesia from all the anaesthetists participating in the study will be observed.

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Patient Information Sheet. Final Version 1.0. Date 17th April 2013

Appendix 16

Two experts will objectively assess all the performances of ultrasound-guided regional anaesthesia independently and will not influence the participating anaesthetist in any way. This assessment will be in real time during the performance of the ultrasound-guided regional anaesthesia and will begin with the initial preparation and equipment setup and end at completion of the procedure.

Why have I been invited?

You are being invited to take part because you are having an operation in one of the theatres in Nottingham University Hospitals NHS Trust where we would like to undertake these observations of anaesthetists performing ultrasound-guided regional anaesthesia. We are inviting 50 participants like you to take part.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. This would not affect the standard of care you receive.

What will happen to me if I take part?

As the study will take place during routine operating lists, therefore, we are only observing an already happening, routine clinical procedure. All anaesthetists are fully competent to carry out the procedure, it is only for the purpose of testing the assessment tool, we will observe their actions that is how they are using ultra-sound to perform your regional anaesthesia.

Two researchers will observe the actions of the anaesthetist while they are using ultra-sound to perform your regional anaesthesia. They will discuss with the anaesthetic team beforehand where the best place is to observe without interfering with the flow / process of safe anaesthesia.

For each individual case, the observations will stop when you leave the anaesthetic room. Each observation will take the length of the Ultra-sound guided regional anaesthesia procedure, which is around **20 minutes** on average. The whole study is taking place over 3 months.

Your name will not be recorded as part of the study and there is no way that any observations could be linked to you.

Expenses and payments

You (any patient) will not be paid to participate in the study.

What are the possible disadvantages and risks of taking part?

We do not anticipate any risk or disadvantage to you from taking part.

What are the possible benefits of taking part?

We cannot promise the study will help you but the information we get from this study may help to determine the validity and reliability of the assessment tool for the performance of ultrasound-guided regional anaesthesia.

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Patient Information Sheet. Final Version 1.0. Date 17th April 2013

What happens when the research study stops?

The study will only be stopped prematurely if we fail to recruit participants.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers contact details are given at the end of this information sheet. If you remain unhappy and wish to complain formally, you can do this by contacting patient advice and liaison service (PALS) at Nottingham university hospitals NHS trust. The contact details are given at the end of this information sheet.

Will my taking part in the study be kept confidential?

We will follow ethical and legal practice and all information about you will be handled in confidence.

If you join the study, some parts of the data collected for the study will be looked at by authorised persons from the University of Nottingham who are organising the research. They may also be looked at by authorised people to check that the study is being carried out correctly. All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.

All information, which is collected, about you during the course of the research will be kept **strictly confidential**, stored in a secure and locked office, and on a password protected database. Any information about you that leaves the hospital will have your name and address removed (anonymised) and a unique code will be used so that you cannot be recognised from it.

All research data will be kept securely for 7 years. After this time your data will be disposed of securely. During this time all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team will have access to your personal data.

What will happen if I don't want to carry on with the study?

Your participation is voluntary and you are free to withdraw at any time, without giving any reason, and without your medical care or legal rights being affected. If you withdraw then the information collected so far cannot be erased and this information may still be used in the project analysis.

What will happen to the results of the research study?

The results will be presented as a written project authored by Dr Atif Shafqat in completion of his PhD. The findings of this study may also be presented at national or international meetings and may be published in a medical journal.

Who is organising and funding the research?

This research is organised by the University of Nottingham and is being funded by the institutional funds. There is no external funding for this study.

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Patient Information Sheet. Final Version 1.0. Date 17th April 2013

Appendix 18

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by **Berkshire Research Ethics Committee.**

Further information and contact details

Professor Jonathan Hardman
Chief Investigator
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Tel: 0115 823 1002
Email: J.Hardman@nottingham.ac.uk

Dr Atif Shafqat
PhD. Student Researcher, University of Nottingham
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Tel: 0115 823 1004
Email: msxas4@nottingham.ac.uk

Patient Advice and Liaison Service (PALS)
NUH NHS Trust, c/o PALS,
Freepost, NEA 14614,
Nottingham NG7 1BR
Email: pals@nuh.nhs.uk
Fax: 0115 875 4655
Tel: 0800 052 1195 (City Hospital campus)
     0800 183 0204 (QMC campus)

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Patient Information Sheet. Final Version 1.0. Date 17th April 2013

Appendix 19

The University of Nottingham

Nottingham University Hospitals NHS
NHS Trust

CONSENT FORM - Patient
Final version 1.0. Date 17th April 2013

Name of Patient:

Title of Project:

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Name of Researchers:

Dr. A Shafqat, Dr. V Thanawala, Dr. KM Rafi, Dr. N Bedforth, Dr. R McCahon and Prof. J Hardman.

Please initial all boxes

1. I confirm that I have read and understand the information sheet (<u>Final Version number **1.0** dated the **17th April 2013**</u>) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that information may still be used in the project analysis.

3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential.

4. I agree to take part in the above study.

Name of Participant   **Date**   **Signature**

Name of Researcher   **Date**   **Signature**
taking consent.

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Patient Consent Form. Final Version 1.0. Date 17th April 2013
When completed: - 1 copy for Patient, 1 copy for Patient's Medical Notes, 1 (original) copy for researcher site file.

Page 1 of 1

Appendix 20

The University of Nottingham

DEMOGRAPHIC DATA COLLECTION FORM
Final version 1.0. Date 17th April 2013

Participant's Study Ref: _____

Type of Block Performing: _____

Title of Project:
Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Please complete 3 questions regarding your UGRA practice below;

Total number of ultrasound-guided nerve blocks performed in the preceding Year.	
Number of ultrasound-guided nerve blocks performed in the preceding Month.	
A self-assessment of your UGRA expertise: **Novice, Competent, or Expert.**	

Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia. Demographic Data Collection Form. Final Version 1.0. Date 17th April 2013

Appendix 21

Construct validity and reliability of an Objective structured assessment tool for performance of UGRA.
Assessment Checklist Data Collection Form. Final Version 1.0. Date 17th April 2013

Assessment Checklist Data Collection Form*

Participant's Study Ref:

Title of Project:
Construct validity and reliability of an Objective structured assessment tool for performance of ultrasound-guided regional anaesthesia.

Tasks	Not Performed 0	Performed Poorly 1	Performed Well 2
1) Proper positioning of patient			
2) Correct placement of ultrasound machine relative to patient to allow easy visualization of both			
3) Choice of correct transducer			
4) Correct depth, gain and focal zone choices			
5) Holds the probe appropriately (3 fingers holding the probe and 1 finger touching the patient)			
6) Knowledge or confirmation of screen orientation (ie, which side of probe corresponds to which side of screen)			
7) Scanning of anatomy and proper identification of target			
8) Use of Doppler to rule out vascular structures (if applicable)			
9) Appropriate needle alignment			
10) Maintainance of needle tip image during advancement of needle			
11) Efficiency of regaining needle tip position image (PART Maneuver)			
12) Recognition of proper nerve stimulation at appropriate levels (if nerve stimulation used)			
13) Ensure that current is not <0.2 mA (if nerve stimulation used)			
14) Ask for initial aspiration to rule out intravascular injection			
15) Visualization of needle tip before injection			
16) Ask for 1- to 2-mL intial injection to rule out intraneural and intravascular injection			
17) Ask patient or at least look for signs of pain/discomfort			
18) Ask for proper aspiration every 5-mL increments injection			
19) Recognition of proper needle tip position			
20) Perform appropriate needle tip adjustments			
21) Assessment of ease of injection (high pressure)			
22) Recognition of correct local anaesthetic spread in relation to nerve			

*Ref: Cheung JJ, Chen EW, Darani R, et al. The creation of an objective assessment tool for ultrasound-guided regional anesthesia using the Delphi method. Reg Anesth Pain Med. 2012 May-Jun;37(3):329-33.

Appendix 22

Construct validity and reliability of an Objective structured assessment tool for performance of UGRA. Global Rating Scale Data Collection Form. Final Version 1.0. Date 17th April 2013

Global Rating Scale Data Collection Form*		Participant's Study Ref:		

Score

	1	2	3	4	5
Preparation for procedure (eg, monitors, IV access, U/S machine)	Did not organize equipment well. Has to stop procedure frequently to prepare them.		Equipment generally organized. Occasionally has to stop and prepare them.		All equipment neatly organized, prepared and ready for use.
	[ ]	[ ]	[ ]	[ ]	[ ]
Patient Interaction	Little to no rapport established; Patient is unaware of procedures. No sedation is provided.		Rapport is generally established; patient is aware and informed of most procedures. Patient anxiety is alleviated adequately with a sedative(s).		Strong rapport is established and maintained throughout procedure. Patient is well informed of all relevant procedures. Patient anxiety alleviated through sedation and verbal comforting.
	[ ]	[ ]	[ ]	[ ]	[ ]
Asepsis (eg, use of sterile gloves, proper patient draping, probe sterility, cleansing of skin before infilteration, use of op site)	Practice of proper aseptic technique not generally apparent. Many errors in aseptic technique made throughout procedure.		Generally practices proper aseptic technique. Occasional errors in aseptic technique made during procedure.		Excellently demonstrates proper aseptic technique. Few or no errors in aseptic technique made during procedure.
	[ ]	[ ]	[ ]	[ ]	[ ]
Respect for Tissue	Frequently uses unnecessary force on tissue or causes damage.		Carefully handles tissue but occasionally causes unintentional damage.		Consistently handles tissues appropriately with minimal damage.
	[ ]	[ ]	[ ]	[ ]	[ ]
Time and motion	Many unnecessary movements.		Efficient time/motion but some unnecessary movements.		Clear economy of movements. Maximum efficiency.
	[ ]	[ ]	[ ]	[ ]	[ ]
Instrument handling	Repeatedly makes tentative and awkward movements.		Competent with instruments but occasionally makes awkward or stiff movements.		Fluid movements with instruments and no awkwardness.
	[ ]	[ ]	[ ]	[ ]	[ ]
Flow of procedure	Frequently stops procedure and seems unsure of next move.		Demonstrates some forward planning with reasonable progression of procedure.		Obviously planned course of procedure, with effortless flow from one move to the next.
	[ ]	[ ]	[ ]	[ ]	[ ]
Knowledge of procedure	Deficient knowledge		Knows all important steps of procedure.		Demonstrates familarity with all aspects of the procedure.
	[ ]	[ ]	[ ]	[ ]	[ ]
Overall performance	Very poor		Competent		Clearly superior
Overall, should the candidate; Pass / Fail	[ ]	[ ]	[ ]	[ ]	[ ]

*Ref: Cheung JJ, Chen EW, Darani R, et al. The creation of an objective assessment tool for ultrasound-guided regional anesthesia using the Delphi method. Reg Anesth Pain Med. 2012 May-Jun;37(3):329-33.

CHAPTER – 4

The Impact of Video games on Ultrasound – Guided

Regional Anaesthesia (UGRA) skills

INTRODUCTION

There has been much emphasis on the identification and evaluation of the innate aptitudes of surgical[166,168,232] and anaesthetic trainees[147], since these capabilities have the potential to predict technical expertise of procedural performances. It is believed that technical expertise is influenced by innate or inborn psychomotor and visuospatial abilities.[205] Technical expertise is compared to a motor skill, in which an excellent technique is the outcome of an exclusive blend of innate abilities and practice of a certain task. Expertise differs from ability. Ability is deemed as a set of innate characteristics that decide potential for a given activity. Abilities are present since birth and are comparatively non-customizable despite of any contribution in the form of further practice, training or experience and remain relatively permanent throughout life.[233]

Psychomotor ability is the capability to execute bodily motor movements with precision, coordination, or strength.[234] Expert procedural skillfulness is the product of a precise fusion of both innate psychomotor ability and practice or experience of a particular task.

Visuospatial ability is the aptitude to create, retain, retrieve, and transform well-structured visual images. It is considered to be fundamental in the analysis of images, which are produced by X-ray, computed tomography (CT), magnetic resonance imaging (MRI), or ultrasound.[234] During analysis of such images, medical staffs have to depict mentally and deduce the patients' three-dimensional (3D) anatomy from a two-dimensional (2D) image.[235]

Ultrasound-guided needle placement is a technical procedure that is evolving into a universal skill for medical professionals in a variety of specialities.[236] Medical disciplines including critical care, internal medicine, obstetrics and interventional radiology routinely perform interventional ultrasound-guided procedures.[121,237-240] Ultrasound–guided regional anesthesia (UGRA) is an invasive procedure that involves complex motor skills[241] and requires manual dexterity, hand-eye coordination and a working knowledge of sono-anatomy.[147]

UGRA has many parallels with laparoscopic surgery, particularly with respect to the complex motor skills required.[146] At a basic level, it is possible that laparoscopy and UGRA are similar with respect to the interaction of the operator's hands and eyes with the ultrasound probe/laparoscope, the patient and the screen. The strive for overcoming the technical challenges associated with laparoscopic surgery and UGRA in particular has increased the significance of the physicians' visuospatial and psychomotor abilities. Earlier studies have shown that diversity in innate abilities can predict the rate of skill acquirement.[159,242]

Video games have become an integral part of our daily lives in today's society. In 2002, the total sales in the video game industry came up to $ 10 billion in the United States.[243] The average age of a video gamer is 29 years old.[107] According to Gentile DA et al. the average playing time for 94 % of adolescents is nine hours per week and for adolescent boys this number is even more at 13 hours per week respectively.[244]

Negative correlations with video game playing include lower grades in school[244], childhood obesity[245,246], muscular and skeletal disorders[247,248], epileptic seizures[249] and decrease positive pro-social behaviours.[244] Others also include increases in blood pressure, heart rate and stress hormones.[250,251] However, there are a number of positive benefits of video games. Studies have shown that video game players have superior visuospatial ability[252] and spatial resolution.[253] They have better reaction time[254] and enhanced eye-hand coordination[255] and also demonstrate improved performance on neuropsychological tests[256] and mental rotation.[257]

The influence of nonsurgical aptitude and skillfulness, such as video gaming capability on laparoscopic surgical expertise is still uncertain. For many years there has been an association between video games and surgical aptitude. Game consoles imply the handling of three-dimensional (3-D) objects via a two-dimensional (2-D) screen, which need dexterity of the wrist and hand. This expertise seems analogous to surgical and regional anaesthetic techniques such as laparoscopy and ultrasound-guided regional anaesthesia (UGRA), respectively.

Prior evidence has shown that video games are useful for both visual attention and hand-eye coordination. In a study Green CS and Bavelier D[258] have evaluated the effects of video games on both visual attention and spatial distribution. It was concluded that video gaming was positively correlated to visual attention processing. Similarly, Rosser JC et al.[259] compared surgeons on a laparoscopic training course. They found that the surgeons who had participated in video game activity for more than three hours per week in the past were 27 % faster in completion of the course tasks, made 37 % fewer errors and scored overall 42 % better than surgeons who had

never been involved in any sort of video game activity. Thus, concluded that video game skills correlate with laparoscopic skills.

Other studies[260,261] have recommended that younger surgeons undergo rapid and effortless skills acquisition in laparoscopic surgery than their senior contemporaries. The authors believed that this might be due to the fact that the younger surgeons have been exposed to video games at an early age and thus have had more familiarity and practice with screen-mediated task completion. In addition to that regular involvement in video gaming activities predisposes the subject to be more competent at video-endoscopic surgical tasks.[201,260]

Grantcharov TP et al.[201] examined the impact of video games on virtual reality laparoscopy. The outcome revealed that surgeons with video game experience performed better in terms of number of errors and time completion of the task and this difference remained significant after adjustment for gender and dominant hand.

Recently, there has been rapid growth in studies[201,234,262] suggesting that gamers have better laparoscopic skills and that video games could be used as a cost-effective and fun way to train basic laparoscopic skills in surgical residents. However, to date within regional anaesthesia none of the studies have evaluated the impact of video games on UGRA skills. Therefore the aim of our study is to explore whether playing video games will predict psychomotor performance of ultrasound-guided needling task or could impact on tests of visuospatial ability and also examine the correlation between these innate abilities.

METHODS

The study was reviewed and approved by the University of Nottingham Medical School Research Ethics Committee.

Design

This single centered, prospective, blinded observational study was conducted at the University Department of Anesthesia, Queen's Medical Centre, Nottingham University Hospitals NHS Trust, Nottingham UK.

Subjects

Participants from whom written informed consent has been obtained were included in the study. Whereas subjects with previous experience of ultrasound scanning or performing regional anesthesia were excluded from the study. Participants were free to withdraw from the study at any point in time.

Recruitment

Medical students were invited to participate in the study via a poster [appendix 23] to be circulated through e-mail, and exhibited on the Networked Learning Environment (NLE), by and with permission of the Medical Student Undergraduate Co-ordinator (University of Nottingham). A participant information sheet [appendix 24-27] was forwarded to those who expressed an interest, via e-mail or in person. Candidates wishing to participate were invited to attend and written informed consent [appendix 28] was gained on the day of the study, including consent for video recording.

The study involved three elements [figure 13]. The registered medical students were asked to undergo and complete all three elements of the study. Participants' characteristics were concealed throughout the study, and their assessments were hidden from view within individual folders.
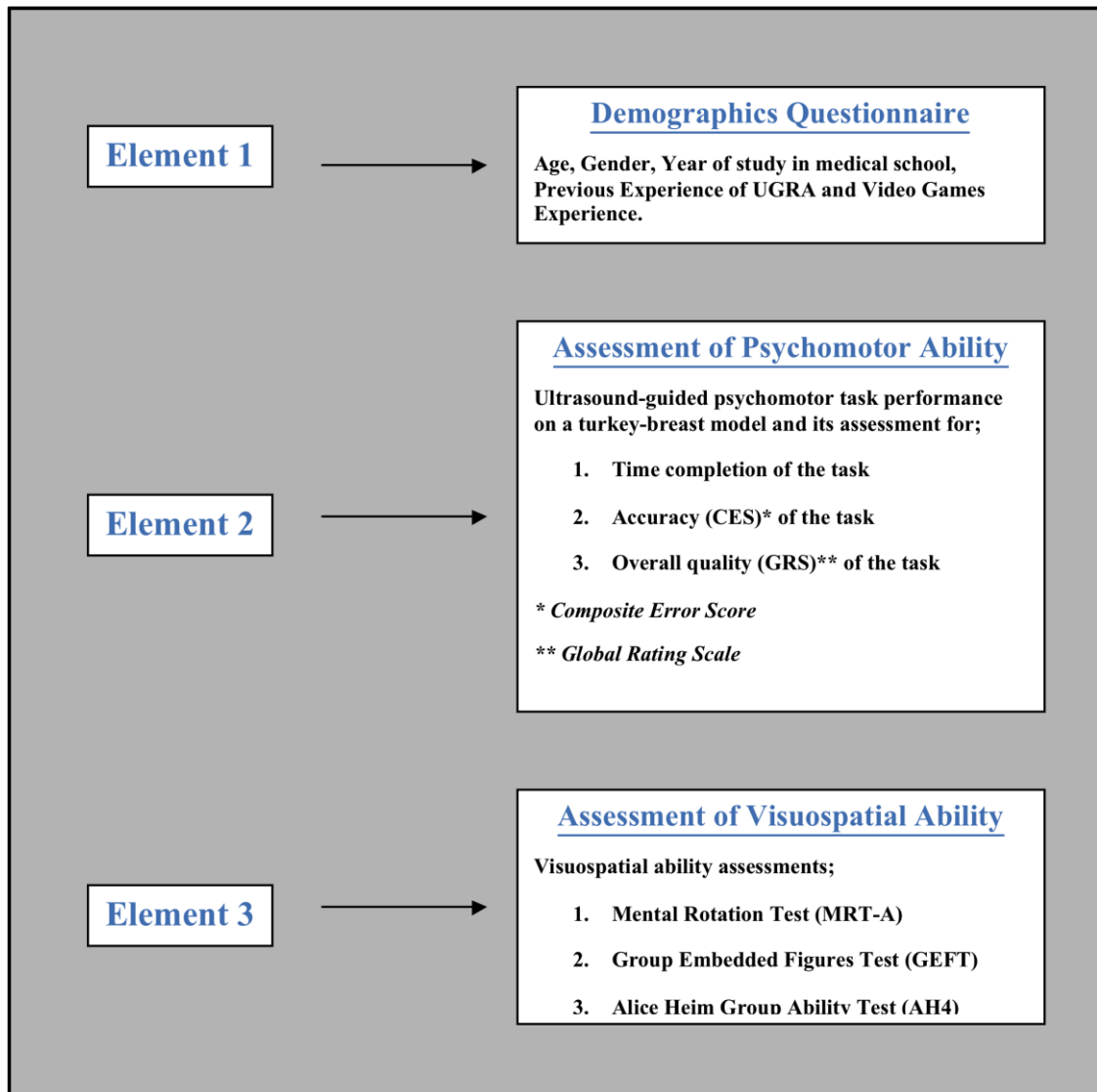

Figure 13 Study Configuration

Study Elements

1. Demographic questionnaire

2. Assessment of Psychomotor ability

3. Assessment of Visuospatial ability

1.   Demographic Questionnaire

The first element of the study entailed completion of a demographic questionnaire. All the participants were asked to complete a questionnaire, providing some background factors like video game experience, previous experience of regional anaesthesia and other demographic information including age, gender and year of study in medical school.

Video game experience was defined as one of the two options: *'occasionally or daily'*. Not having video game experience was defined as *'never played video games'*. Therefore the participants who have previous experience with video games were categorise as **'Gamers'** and those who have no experience at all with video games were categorise as **'Non – Gamers'**.

2.   Assessment of Psychomotor Ability

Element two of the study comprised of evaluation of the psychomotor ability, which was assessed by an ultrasound-guided needling task performance on a turkey-breast model.[192,193] The participants were shown an 11- minute video of expert performance of ultrasound-guided needle advancement, which was mapped to specific learning objectives. The participants had a total of 30 minutes to watch and review the video. They were then asked to complete the needling task on a turkey-breast model, using a standard linear ultrasound probe (38-mm high-frequency linear array transducer; HFL38X 13-6 MHz, Sonosite Limited, Hitchen, UK). In order to increase realism, the turkey-breast was inserted into the groin recess of a Laerdal® IV-Torso manikin (Laerdal Medical Limited, Orpington, Kent, UK). The learning objectives of the ultrasound-guided needling task are summarized in table 8.

Table 8 Learning Objectives of the UGRA task

Objectives of the Ultrasound-guided needling task

    i.   Switch on the ultrasound machine (S-Series, Sonosite Ltd, Hitchen, UK)

    ii.   Correctly orientate the ultrasound probe (linear, 38mm) in relation to the display on the screen.

    iii.   Ensure adequate application of conducting gel to enhance ultrasound transmission and picture quality.

    iv.   Locate and identify the target (olive) within the turkey breast.

    v.   Adjust the gain function to improve the quality of the image by altering brightness of the picture.

    vi.   Alter the depth of the image to obtain a suitable image of the target.

    vii.   Using an in-plane approach insert a 50mm Stimuplex® A needle (B Braun, Melsungen, Germany) into the turkey breast and aim to place the needle tip at the 12 o'clock position, as indicated by the attending assessors, above the upper edge of the target, without piercing the target.

Two anaesthetists experienced in performing ultrasound-guided regional anaesthesia assessed the participants' performance of the needling task independently. The participants received no help or feedback before or during the task. The task performance was evaluated in terms of (i) 'time of completion', (ii) the 'accuracy' and (iii) 'overall quality'. The 'accuracy' was assessed by composite error score (CES), while the 'overall quality' was calculated by global rating scale (GRS).

CES[127,193,194] [appendix 29] and GRS[106,117,146] [appendix 30] are two previously validated assessment tools. The assessors have undergone specific training and practice in the use of these tools. The CES used in our study was similar to the one used by Davies T et al.[194] The CES was calculated by adding together the measurements of 'total number of errors', 'number of needle passes' and 'image quality scale' for each of the participant's task performance. 'Number of needle passes' is defined as, *"a new puncture of the turkey breast skin or the withdrawal of the needle towards the exterior of the turkey breast skin".* This definition along with the classification of 'image quality scale' is similar to the one used by Sites BD et al.[193] A lower composite error score is associated with better accuracy and task performance.

The GRS was a modified version of that used in the original study by Martin JA et al.[120] The GRS consisted of seven items each rated on a five – point scale. The GRS predominantly assessed more general behaviors and the overall performance of the participants.

A video recording of the participant's performance was made by a tripod-mounted camcorder focusing the whole task including the ultrasound image and will only be used for reference purposes, in circumstances if there were any uncertainties concerning the participant's performance.

3. Assessment of Visuospatial Ability

The third and the final element of the study consisted of visuospatial ability assessments. Participants will undergo following assessments of visuospatial ability:

A) Mental rotation test (MRT)

B) Group Embedded Figures test (GEFT)

C) Alice-Heim Group Ability test (AH4)

Assessments of visuospatial ability were paper-based and administered under strict examination conditions in the University Department of Anaesthesia, Queen's Medical Centre, Nottingham. Brief descriptions of all the visuospatial assessments are provided below.

Visuospatial Assessments

The visuospatial assessments consisted of the Mental Rotation Test (MRT), Group Embedded Figures Test (GEFT) and Alice Heim Group Ability Test (AH4).

There are four different variations, based on the original Vandenberg & Kuse (1978) mental rotation test figures, which in turn are based on figures provided by Shepard & Metzler.[183] These include MRT–A, MRT–B, MRT–C and MRT–D respectively.

We used MRT–A, which consists of 24 problem figures. Each problem task has a target figure on the left and four stimulus figures on the right. Two of these stimulus figures are rotated versions of the target figure and two of the stimulus figures cannot be matched to the target figure. The aim is to mentally rotate the figures around the vertical axis to find the two correct rotated versions of the target figure. One point is given if both of the stimulus figures that match the target figure are identified correctly. No point is given for a single correct answer. This means that the maximum score obtainable is 24 points. Participants were given four minutes to complete the first set of 12 problem tasks, followed by a one-minute break before completing the second set of 12 problem tasks in the next four minutes.

B) *The Group Embedded Figures Test (GEFT)*[184,185]

The GEFT was designed to provide an adaptation of the original individually administered Embedded Figures Test (EFT), which would make possible group testing. The GEFT has been modeled as closely as possible on the individually administered EFT with respect to mode of presentation and format. It contains 18 complex figures, 17 out of which were taken from the EFT. The aim is to find a previously seen simple figure within a larger complex figure, which has been

structured in a way to obscure or embed the simple figure. The participants were required to identify and outline accurately a simple shape embedded in a complex figure. The participants were given 10 minutes to complete all of the 18 problems. The score is the total number of simple forms correctly traced within a complex figure. Omitted items are scored as incorrect. In order to receive the credit for an item, all lines of the simple form must be traced; the subject has added no extra lines and all incorrect lines have been erased. In this way the maximum score a subject could attain is 18.

C) *Alice Heim Group Ability Test (AH4)*[186]

The AH4 incorporate 2 parts. Part-1 consists of 65 questions that have a verbal or numerical bias while Part-2 comprises of 65 questions, which have a diagrammatic bias. Problems are in multiple-choice form, the number of proffered answers in every case being five.

In our study we used AH4 to only assess diagrammatic bias in the form of spatial reasoning skills, which is the ability to visualize, mentally rotate and manipulate two-dimensional or three-dimensional shapes or patterns. As mentioned above the test consists of 65 questions and the time limit for completing the test is 10 minutes, exclusive of the preliminary examples. One mark is scored for each correct answer. Therefore the maximum score that could be achieved is 65.

Blinding

Ultrasound-guided needling task assessors were not allowed to observe the assessments of visuospatial ability. Data interpretation was also blinded. Participants were also blinded to the study hypothesis and their test scores.

Primary outcome:

1. Total time to complete the psychomotor task of 'gamers' and 'non – gamers'

2. Composite error scores (CES) of 'gamers' and 'non – gamers'

3. Global rating scale (GRS) of 'gamers' and 'non – gamers'

4. Visuospatial tests scores of 'gamers' and 'non – gamers'

Secondary outcome:

1. Correlation between tests of psychomotor and visuospatial abilities respectively

2. Reliability of the CES and GRS

Statistical Analysis

The statistical analysis was performed using STATA/IC version 10.0 (StataCorp, Texas, USA). Descriptive statistics for demographic and outcome measure data were calculated. Shapiro-Wilk and Skewness / Kurtosis tests were used to determine the normality. CES data follow a non-normal distribution, therefore presented as median (IQR).

Independent samples t-tests was used to compare the means of a normally distributed outcome variable. The Wilcoxin-Mann-Whitney test analysed non-normally distributed outcome variable. Linear regression analysis was used to assess the significant relationships. Count data was analysed using negative binomial regression model. Relationships between psychomotor and visuospatial abilities were examined using Spearman's rank correlation coefficient 'ρ' (Rho). To achieve a study power of 0.8 (α = 0.05), we calculated that we would need to recruit 60 participants for this model with an assumed moderate effect size[195] (r = 0.3 – 0.5).

Reliability of the assessment tools i.e. CES and GRS, was evaluated using Intra-class correlation coefficient (ICC), Cronbach's alpha coefficient and standard error of the measurement as a percentage of the mean (SEM%).[196,197] In all cases, we considered P – values less than 0.05 (two-tailed) to indicate statistical significance.

RESULTS

All the participants of the study completed each of the three elements of the study and as a result there was no missing data. In total, 17 out of 60 participants played video games ('gamers') [table 9] according to the criteria defined in the methods section. The average age of 'gamers' and 'non – gamers' were 23.2 and 23.3 years respectively. Other demographics of the participants are presented below in table 9.

Table 9 Demographics

Participant Demographics		
	Gamers	**Non – Gamers**
Total N (%)	17 (28.3)	43 (71.7)
Gender		
Male (n, %)	14 (82.3)	16 (37.2)
Female (n, %)	3 (17.6)	27 (62.8)
Age in years, Mean (SD)	23.2 (3.7)	23.3 (4.6)
Year of study in medical school*		
Year 1 (n, %)	0 (0.0)	8 (18.6)
Year 2 (n, %)	8 (47.0)	13 (30.2)
Year 3 (n, %)	2 (11.8)	11 (25.6)
Year 4 (n, %)	3 (17.7)	8 (18.6)
Year 5 (n, %)	1 (5.9)	3 (7.0)

* Year of study in medical school (Video Gamers only);

3 (17.6%) Missing values

Psychomotor Ability

i) Time completion of the Psychomotor task

There was a significant difference between 'gamers' and 'non – gamers' in time to complete the psychomotor task (P = 0.02) [table 10, figure 14], such that 'gamers' were approximately two minutes faster than 'non – gamers'.

A linear regression model showed that videogame experience explained significantly the variation in time completion of the psychomotor task ($F_{1,58}$ = 5.4, P = 0.02). An inverse relationship was observed between videogame experience and the time completion of the psychomotor task. The regression equation showed that we would expect 'gamers' to finish the task by 1.98 minutes earlier than 'non – gamers', i.e. 'gamers' had completed the psychomotor task quicker and faster as compared to 'non – gamers'.
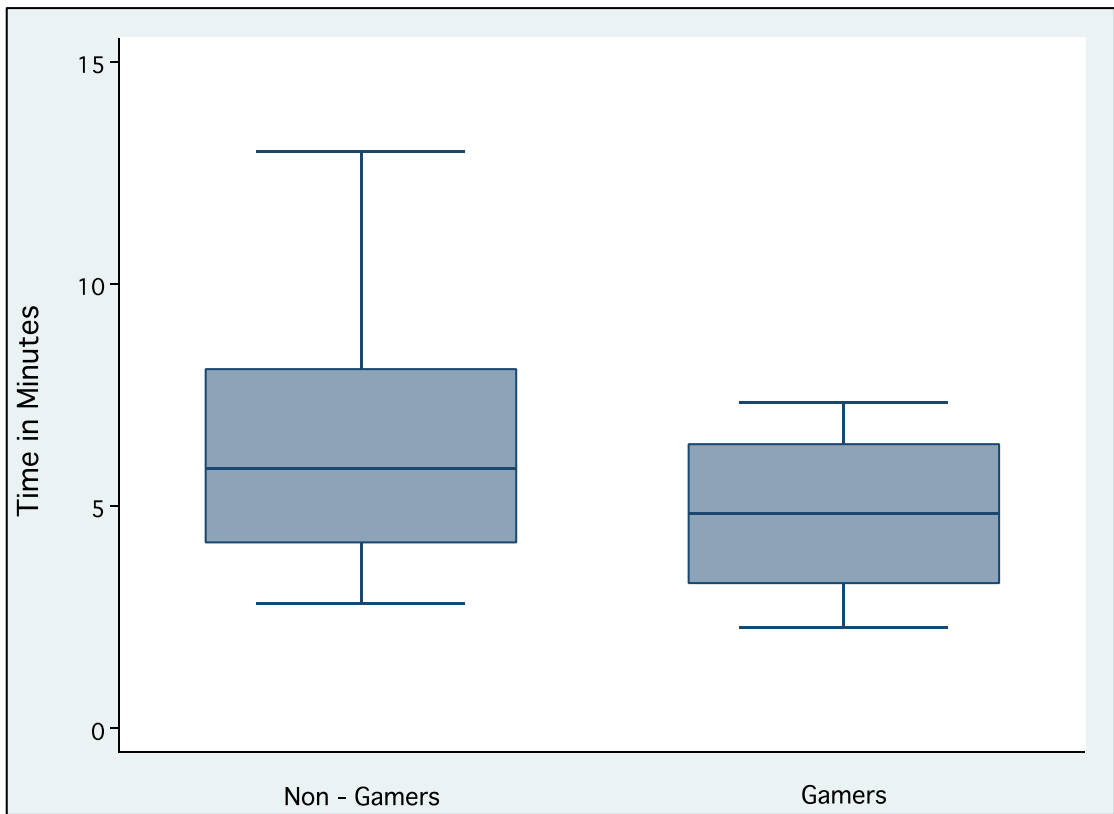
Figure 14 'Time Completion' of the Psychomotor task for 'Non - Gamers' and 'Gamers'

ii) Accuracy (CES) of the Psychomotor task

A significant difference was found between 'gamers' and 'non – gamers' for CES (P = 0.04) [table 10, figure 15].

The negative binomial regression model for CES showed that videogame experience explained significantly the variation in CES (P = 0.01). An inverse relationship was seen between the videogame experience and CES. The regression equation showed that 'gamers' made 0.65 fewer errors than 'non – gamers' i.e. the accuracy of the psychomotor task was better for 'gamers' than 'non – gamers'.
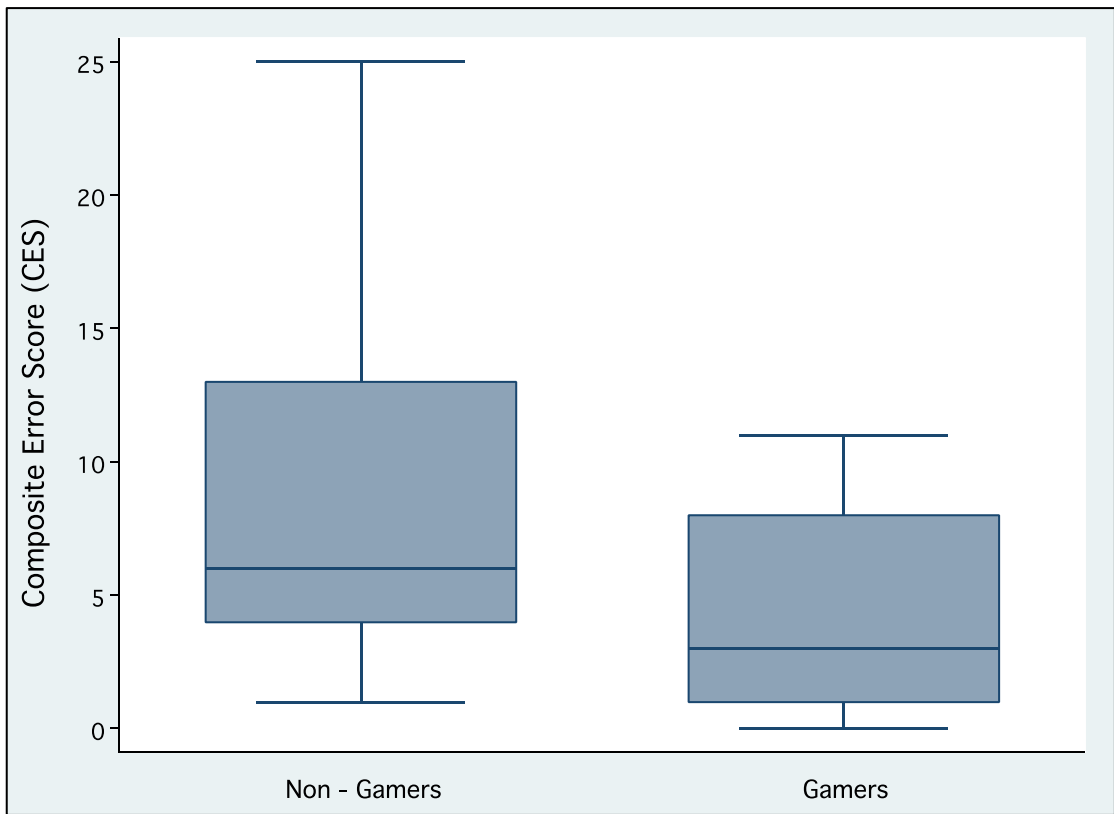
Figure 15 Accuracy (CES) scores of the Psychomotor task for 'Non - Gamers' and 'Gamers'
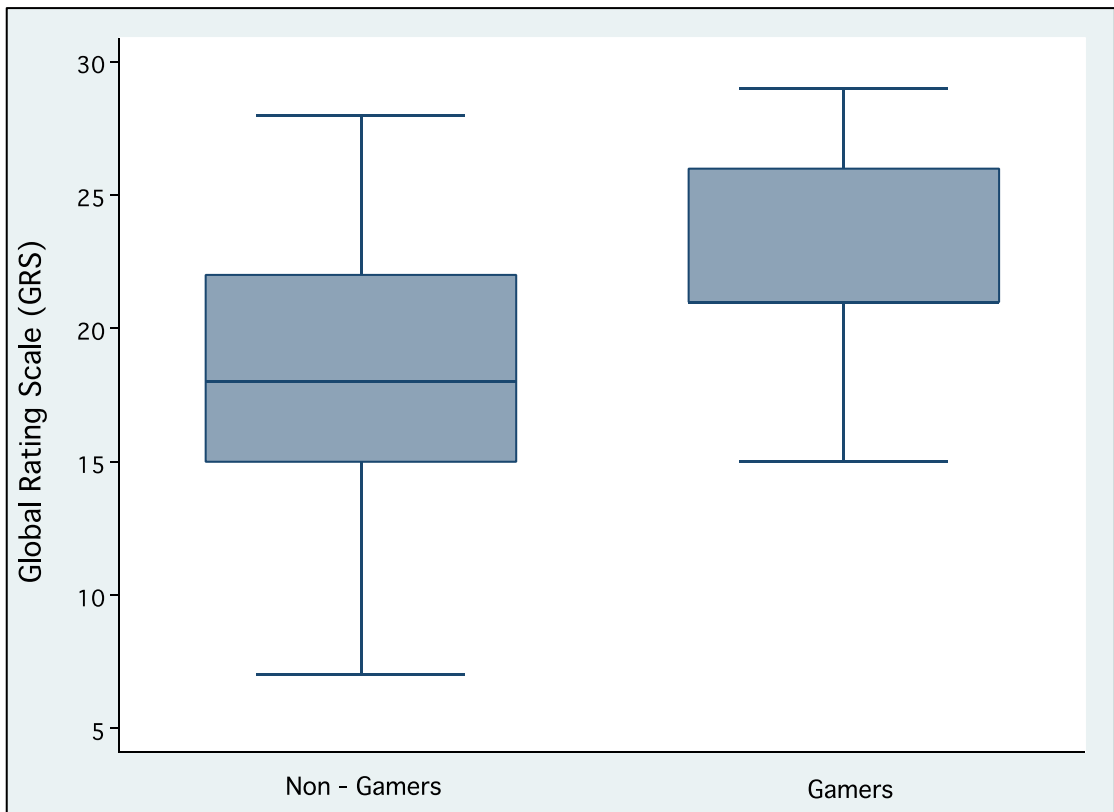
Table 10 Psychomotor and Visuospatial ability assessment scores for 'Non - Gamers' and 'Gamers'

	Non - Gamers	Gamers	P-value	Tests
Psychomotor Ability:				
Time completion of the psychomotor task in mins. Median (IQR)	5.8 (3.9)	4.8 (3.1)	0.02	Mann-Whitney
CES Median (IQR)	6.0 (9)	3.0 (7)	0.04	Mann-Whitney
GRS Mean (SD)	18.3 (5.4)	23.5 (6.0)	0.007	Mann-Whitney
Visuospatial Ability:				
MRT Mean (SD)	12.8 (5.2)	16.9 (4.6)	0.006	Mann-Whitney
AH4 Median (IQR)	58 (12)	62 (3)	0.09	Mann-Whitney
GEFT Median (IQR)	17 (3)	17 (3)	0.7	Mann-Whitney

Mean and SD (standard deviation) used for continuous variables with a normal distribution and median and IQR (Interquartile range) for continuous variables not normally distributed. Significance at $P < 0.05$. CES – Composite Error Score; GRS – Global Rating Scale; MRT – Mental Rotation Test; AH4 – Alice Heim Group Ability Test; GEFT – The Group Embedded Figures Test.

iii) Overall quality (GRS) of the Psychomotor Task

GRS was significantly higher for 'gamers' as compared to 'non – gamers' (P = 0.007) [table 10, figure 16].

A linear regression model was used to further explore the relationship. It showed that videogame experience explained significantly the variation in the GRS ($F_{1,\,58}$ = 10.37, P = 0.0021). A direct relationship was observed between videogame experience and GRS. The regression equation shows that we would expect 'gamers' to increase the GRS by 5.17 unit when compared to 'non – gamers' i.e. the overall quality of the psychomotor task performance was better in 'gamers' than 'non – gamers'.

Figure 16 Overall quality (GRS) scores of the Psychomotor task for 'Non - Gamers' and 'Gamers'

Visuospatial Ability

I. MRT

The MRT scores were significantly higher for 'gamers' as compared to 'non – gamers' (P = 0.006) [table 10, figure 17].

A linear regression model was used to examine the relationship. It showed that videogame experience explained significantly the variation in the MRT score ($F_{1, 58}$ = 7.99, P = 0.006). A direct relationship was observed between videogame experience and MRT. The regression equation showed that we would expect 'gamers' to increase the MRT scores by 4.09 unit when compared to 'non – gamers'.
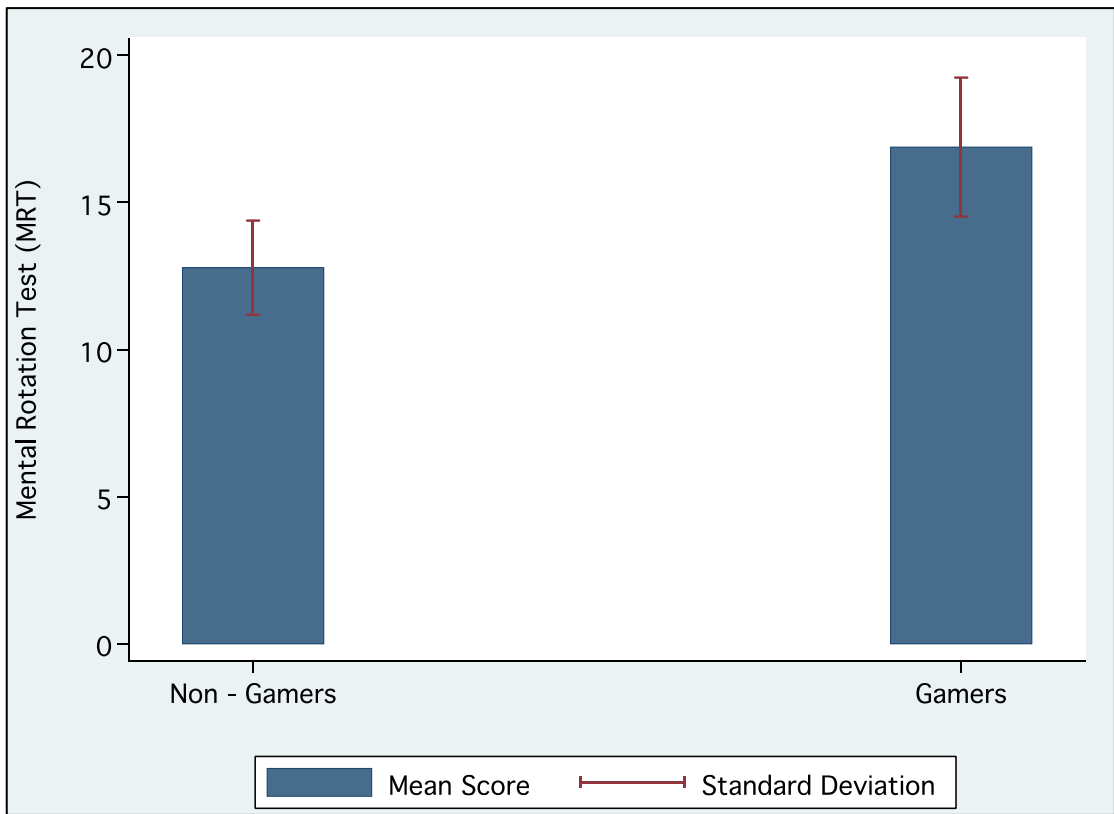
Figure 17 MRT scores for 'Non - Gamers' and 'Gamers'

II. AH-4 & GEFT

No significant differences were found between 'gamers' and 'non – gamers' for the AH-4 scores (P = 0.09) and the GEFT scores (P = 0.7) respectively [table 10, figures 18 & 19].

Similarly the linear regression model also revealed no significant difference between 'gamers' and 'non – gamers' for the AH-4 ($F_{1, 58}$ = 3.87, P = 0.05) and the GEFT ($F_{1, 58}$ = 0.32, P = 0.573), respectively.
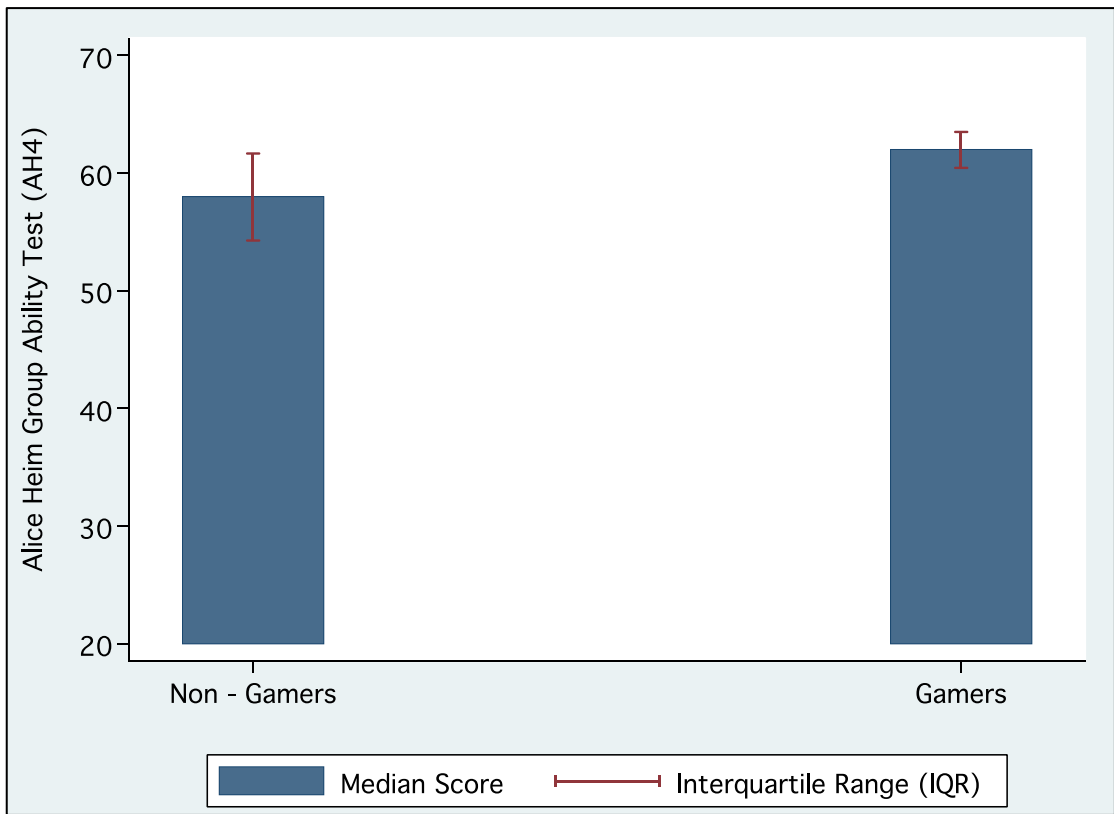
Figure 18 AH4 scores for 'Non - Gamers' and 'Gamers'
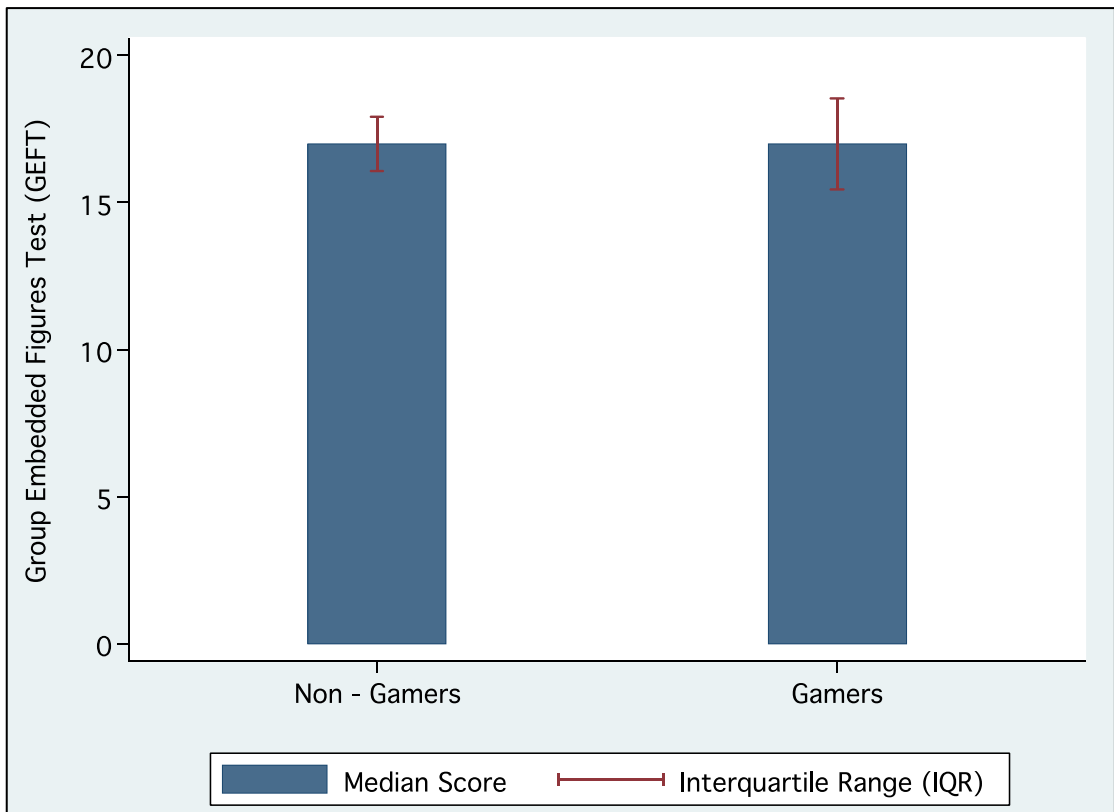
Figure 19 GEFT scores for 'Non - Gamers' and 'Gamers'

Relationship between Psychomotor and Visuospatial Abilities

Of the three psychomotor abilities 'time completion' correlated significantly with CES, GRS and MRT respectively (table 11). There was a positive correlation between 'time completion' versus CES (ρ +0.69, P <0.001) i.e. the participants who took more time to complete the ultrasound-guided needling task made more errors. Similarly, there was negative correlation between 'time completion' versus GRS (ρ −0.71, P <0.001) and 'time completion' versus MRT (ρ −0.40, P =0.001) respectively i.e. the participants who needed more time to complete the needling task scored less in GRS and MRT tests. No significant correlation was found between 'time completion' and remaining visuospatial abilities including AH-4 (ρ −0.07, P =0.56) and GEFT (ρ −0.03, P =0.80) respectively.

There was negative correlation between CES versus GRS (ρ −0.78, P <0.001) and CES versus MRT (ρ −0.54, P <0.001) respectively, signifying that higher error rate was associated with low GRS and MRT scores. However, GRS had a linear correlation with MRT (ρ +0.47, P <0.001), indicating that better overall quality of the task performance is linked with enhanced MRT skills.

Table 11 Correlation (rho) 'ρ' between Psychomotor and Visuospatial ability assessments

Significant relationships between Psychomotor & Visuospatial abilities	Correlation 'ρ' (rho), P - Value
Time to complete the needle task Vs CES	ρ + 0.69, P < 0.001
Time to complete the needle task Vs GRS	ρ − 0.71, P < 0.001
Time to complete the needle task Vs MRT	ρ − 0.40, P = 0.001
CES Vs GRS	ρ − 0.78, P < 0.001
CES Vs MRT	ρ − 0.54, P < 0.001
GRS Vs MRT	ρ + 0.47, P < 0.001

Spearman's correlation coefficient (ρ) of Psychomotor and Visuospatial Abilities. Significance at P < 0.05. CES – Composite Error Score; GRS – Global Rating Scale; MRT – Mental Rotation Test.

Reliability of CES and GRS

The intra-class correlation coefficient and SEM (%) for CES and GRS was 0.97 (15.29 %) and 0.91 (8.53 %) respectively; this demonstrates a high degree of inter-rater agreement. Similarly, Cronbach's alpha coefficient and SEM (%) for CES and GRS was 0.98 (9.49 %) and 0.96 (5.69 %) respectively; this demonstrates a high degree of inter-item consistency.

DISCUSSION

In our study, 'gamers' performed significantly better than 'non – gamers' in the psychomotor performance of an ultrasound-guided needle task. The 'gamers' constantly throughout the study required less time to complete the needle task. In addition to this they also improved the overall quality and the accuracy of the psychomotor performance of an ultrasound-guided needle task. It is the accuracy or in other words decreased errors that will have the most important effect on patient safety.

This is a correlational study and therefore causality cannot be definitely concluded. Some medical students may possibly have enhanced psychomotor abilities, which predispose them to more dexterous activities including video games. Other medical students may well merely love playing video games and as a result have a habit of playing more often, therefore strengthening their motor skilfulness. Similarly, one more hypothesis is that innate dexterous people will be engaged to video games, because top scores will be more gratifying for them.[255] According to Miskry T et al.[263] dexterous non-gamers have demonstrated improved success at a racing game than their colleagues.

In addition, there may be other confounding variables, including interests in pastime such as sports or music. The medical students in our study are novices to ultrasound-guided regional anaesthesia; therefore the possible confounding factor of previous experience in ultrasound-guided regional anaesthesia is removed.

Hypothetically, there should be no impact of video gaming on visuospatial abilities since they are regarded as innate and for that reason should remain unchanged. On the other hand, our study contradicts with this theory. Out of the three visuospatial ability assessments, mental rotation correlated significantly with video games practice. 'Gamers' scored higher in the MRT as compared to 'non – gamers'. The change in visuospatial abilities due to video games practice is further endorsed by the literature.[258,264] Green CS and Bavelier D[258] illustrated that video game practice improves visual attention capacity and spatial distribution. Similarly, Li R et al.[264] confirmed that contrast sensitivity (the ability to detect small increments in shades of gray on a uniform background) augments with video game practice.

It is essential to reflect about the visuospatial assessments we employed. Even though all the assessments were validated tools, but they evaluated only certain elements of innate abilities. For instance, the AH-4 test consists of two components, which can assess both the verbal or numerical bias and diagrammatic bias respectively. We used AH-4 for assessing spatial skills via diagrammatic bias only. However, not all these elements may be applicable to UGRA skills. For that reason, it might be desirable to carry out an analogous study with larger number of participants to explore this point even more or to investigate diverse characteristics of innate abilities.

Video games have become an integral part of global culture. According to DFC intelligence, which is a strategic market research and consulting firm, the global video game industry will reach $100 billion by 2019.[265] Since they are practically low-priced, handy and dependable thus it is reasonable to investigate the positive advantages for training with video games. Video game play has been integrated into

training by industries and organizations in which standard training scenarios are too risky or costly. In fact, the United States military are previously using video games such as *Delta Force 2* (NovaLogic Inc, Calabasas, California) as part of their training tactic.[266] Similarly the US army has also approved the popular violent video game series *Rainbow Six* to train its special operations forces since it is a brilliant technique to teach all the essential phases required to plan and carry out a successful special-operations mission.[267] Hopefully as an analogy, medicine can tap into the training and technology guidance of video games.

There is ample evidence available, which suggest that a positive correlation exists between video gaming and generalised surgical ability including open, laparoscopic, endoscopic and robotic surgery.[268-270]

Prior studies have examined a potential relationship between video gaming and laparoscopic skills using conventional training modules. Rosenberg BH et al.[271] in their study found a significant correlation between video gaming and laparoscopic tasks performed on pigs in a group of 11 students. Badurdeen S et al.[272] in their retrospective cohort study assessed medical students and junior doctors with minimal laparoscopic experience on Nintendo Wii games and laparoscopic simulator tasks. They concluded that prior video game experience positively correlated with laparoscopic scores.

The impact of video games on virtual reality laparoscopic simulator skills has also been investigated. Grantcharov TP et al.[201] were the first to explain a positive relationship between video games experience and virtual reality laparoscopic

simulator score. The study subjects with prior video game experience made significantly less errors on the simulator, even after adjusting for sex and hand dominance. In the same way, Rosser JC et al.[259] confirmed that previous video gaming practice of more than three hours per week correlated with considerable less errors and quicker completion of a certified laparoscopic skills and suturing program for surgeons.

An Irish study by Kennedy AM et al.[234] analysed the psychomotor, perceptual and visuospatial skills of medical students. The outcome revealed that video gaming illustrates a positive association with the psychomotor skills but failed to establish any significant relationship with the perceptual and visuospatial skills respectively. There was a considerable difference in scores between gamers and non – gamers, in terms of time taken to complete the tasks and instrument path length scores (which is a measure of the efficiency by which a target is reached).

Schlickum MK et al.[273] carried out the only randomized controlled trial on this topic in which medical students were randomly assigned to play either a visuospatial video game (the Half Life game) or a Chessmaster game for a period of five weeks until then their laparoscopic simulator skills were compared to the control group. The results revealed positive association between laparoscopic performance and self-reported past video game experience. Both video games group performed considerably better on the laparoscopic simulator.

In another retrospective cohort study by Shane MD et al.[262] 11 medical students and 15 surgical residents were queried about their video games experience. Then they were assessed on the number of trials necessary to achieve competency on a laparoscopic simulator. Those participants who had video games experience of more than three hours per week achieved competency in less time in contrast to the participants with no experience of video games. Thus the amount of training duration in terms of times and repetitions in order to achieve competency in basic skills curriculum of the laparoscopic simulators reduces significantly with video games experience. This was further endorsed via two analogous studies by Stefanidis D et al.[161] and Hogle NJ et al.[274] respectively.

Limitations are associated to any study and therefore needed explicit comment here. This included lack of randomization, absence of a control group and choice of a single institution rather than multiple centres. At present there is no standard criterion to define "*video game experience*". Similarly, there is no validated, model questionnaire to record the "previous or current experience". Rosser JC et al.[259] developed a scale for their study to record 'the amount of video game experience' but did not publish it for further public use.

Similarly, other studies employed absolutely diverse methods for documenting game experience. This included total number of hours of game experience[272], self-reported by the participants as novice, expert and non-gamer[275], or whether an individual had an interest and merely 'likes TV games'.[276] According to Van Dongen KW et al.[277] video games experience was classified as 'an average playing time of at least 10 hours per week'. The study did not succeed in accumulating a group of individuals who

166

played 10 hours per week. On the contrary, the study participants who were the interns played an average of only 1.9 hours per week. A good quality scale is the one, which in addition to report the total hours of game play should also illustrate the dispersion of these hours over a lifetime. The reason for this is that children have more free time to play video games than students and working professionals. This problem was resolved in a study by Schlickum MK et al.[273] in which the participants were allowed to score their game experience on a 7-point Likert scale (1 = never playing, 7 = playing everyday) for current experience and between ages 1 to 6, 7 to 12 and 13 to 18, respectively. Therefore it is unfeasible to evaluate these different methods on an equivalent status.

Simulators have already established their effectiveness in laparoscopy but they are costly, mostly boring and not rapidly available. On the contrary video games are economical, could be used for teaching and are chosen over lectures.[278,279] A spirit of competition develops with other peers, which also increases the enthusiasm to willingly join skills training.[280] Video games may possibly be a component of the training program for basic laparoscopic skills in surgical novices.[281] According to the literature video game exclusively intended to develop basic laparoscopic skills could make virtual reality training cheaper and more enjoyable.[272,273,280,282] However, video games can never be an alternative to true simulators and real operating room experience.

Due to broad and continued charm of video games, it seems logical to investigate their positive effects for education, skill attainment, and maintenance. The results of this study should encourage a discussion regarding how to incorporate video game play as an adjunct to UGRA skills and hence reflect on using video games as another educational tool in curricula suitable for individual trainees. In other words trainees who have superior video gaming skills may possibly use games as one of the element of their training programme, while those trainees who acquire little or no video gaming skills can use other training modalities.[283] However, before implementing all this a comprehensive preliminary evaluation of baseline skill is indispensable.[284]

There is a correlation between video game experience and UGRA skills. However, the predictive significance of these conclusions is unclear. The existing evidence is too weak to absolutely determine the positive outcome of gaming in the improvement of this expertise. Further research is essential to identify the role of video games on UGRA skills training and future trials should be more standardized and sufficiently powered, so stronger evidence on this subject can be attained.

Appendix 24

The University of Nottingham

University of Nottingham
School of Clinical Sciences
Division of Anaesthesia and Intensive Care
Queen's Medical Centre
Derby Road
Nottingham
0115 8231009

Participant Information Sheet

The impact of video games on ultrasound-guided regional anaesthesia (UGRA) skills.

Researchers:

Dr. A Shafqat, Dr. N Bedforth, Dr. R McCahon and Prof. J Hardman.

Invitation paragraph

You have been invited to take part in a research study. Before you decide whether to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends and relatives if you wish to. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you wish to take part or not. If you decide to take part you may keep this leaflet. Thank you for reading this.

Background

Video games have become an integral part of global culture. In today's society they are regarded as one of the essential genre of leisure activity. Video games are renowned for negative correlations associated with it. However, there are a number of positive benefits of video games. For many years there has been an association between video games and surgical aptitude. There is sufficient evidence available, which suggest that a positive correlation exists between video gaming and surgical ability.

We aim to identify whether playing video games will predict psychomotor performance of an ultrasound – guided needling task or could impact on tests of visuospatial ability and also examine the correlation between these innate abilities.

What does the study involve?

The study will be conducted in the University department of Anaesthesia, Queen's Medical Centre Campus, Nottingham University Hospitals NHS Trust, U.K. It is an observational study involving healthy medical students from the University of Nottingham. Potential volunteers wishing to participate will be invited to attend and informed consent will be gained on the day of the study, including consent for video recording.

Appendix 25

The study consists of three elements below:
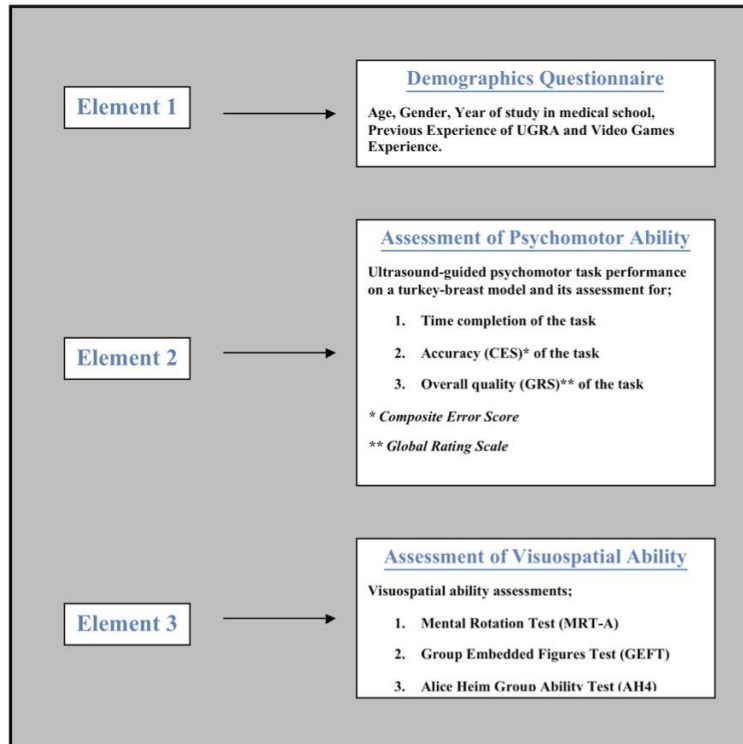
Element 1 → **Demographics Questionnaire**
Age, Gender, Year of study in medical school, Previous Experience of UGRA and Video Games Experience.

Element 2 → **Assessment of Psychomotor Ability**
Ultrasound-guided psychomotor task performance on a turkey-breast model and its assessment for;

1. Time completion of the task
2. Accuracy (CES)* of the task
3. Overall quality (GRS)** of the task

Composite Error Score
**Global Rating Scale*

Element 3 → **Assessment of Visuospatial Ability**
Visuospatial ability assessments;

1. Mental Rotation Test (MRT-A)
2. Group Embedded Figures Test (GEFT)
3. Alice Heim Group Ability Test (AH4)

Assessment of the Ultrasound Task

Two anaesthetists, experienced in performing ultrasound-guided regional anaesthesia, will assess the volunteer's performance of the task independently. The video recording will only be used for reference if there are any uncertainties concerning the volunteer's performance.

171

Appendix 26

Why have you been chosen?
You have been chosen because you are aged over 18 years, you are healthy, and you are able to give informed consent to participate in this study.

Do you have to take part?
It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

What are the possible disadvantages and risks of taking part?
There are no known health issues associated with the use of ultrasound. Possible side effects include accidental needle prick injury during handling. If you feel uncomfortable in any way, you can stop at any time. You will also need to be able to spare a total of 150 minutes of your total time.

What if something goes wrong?

If you should have any complaints of any aspect of the study you should initially contact;

Professor Jonathan Hardman
University Department of Anaesthesia
Queens Medical Centre, Nottingham NG7 2UH
Telephone 0115 8231002
E-mail J.Hardman@nottingham.ac.uk

If you remain unhappy and wish to complain formally, please contact the secretary of the University of Nottingham Medical School Research Ethics Committee, Mrs Louise Sabir, E-mail address: louise.sabir@nottingham.ac.uk

In the unlikely event that you suffer injury to yourself or damage to your property as a result in taking part in this research, the University does have an insurance policy to cover harm arising as a result of the defect in the design of the study. In addition, all medical practitioners taking part in the research have personal medical negligence cover.

Appendix 27

Will my taking part in this study be kept confidential?

All information, which is collected, about you during the course of the research will be kept on a password protected database and is strictly confidential. Any information about you that leaves the research unit will have your name and address removed so that you cannot be recognised from it. All information generated by this study will be archived securely within the University Division of Anaesthesia & Intensive Care, University of Nottingham and destroyed 7 years after the study is completed.

What will happen to the results of the research study?

The results will be presented as a written project authored by Dr Atif Shafqat in completion of his PhD. The findings of this study may also be presented at national or international meetings and may be published in a medical journal.

Who is organising and funding the research?

The study is part of post-graduate research supported by The University of Nottingham. There is no external funding.

Who has reviewed the study?

This study has been reviewed and approved by the University of Nottingham Medical School Ethics Committee.

Contact for Further Information

Dr Atif Shafqat
PhD. Student University of Nottingham
University Department of Anaesthesia
Queens Medical Centre
Nottingham NG7 2UH
Email: msxas4@nottingham.ac.uk

Many thanks for your participation in this study.

Page 4 of 4
The impact of video games on UGRA skills. Participant Information Sheet. V 1.0. Date 10th June 2014

173

Appendix 28

University of Nottingham
School of Clinical Sciences
Division of Anaesthesia and Intensive Care
Queen's Medical Centre
Derby Road
Nottingham
0115 8231009

PARTICIPANT CONSENT FORM

The impact of video games on ultrasound-guided regional anaesthesia (UGRA) skills.

Name of Researchers:
Dr. A Shafqat, Dr. N Bedforth, Dr. R McCahon and Prof. J Hardman.

Study Number:_____

Please write your **initials** in each box

1. I confirm that I have read and understand the information sheet (version number **1.0** dated the **10th June 2014**) for the above study and have had the opportunity to ask questions.

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that information may still be used in the project analysis.

3. I understand that all data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to me taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential.

4. I understand and agree that I may be video-recorded during my participation in the study and that this video recording will only be used for reference by the study investigators.

5. I agree to take part in the above study.

_____   _____   _____
Name of Participant   Date   Signature

_____   _____   _____
Name of Researcher   Date   Signature

When completed: - 1 copy for participant, 1 copy for researcher site file.
Participant Consent Form. Impact of video games on UGRA skills. University of Nottingham. Final Version 1.0

Appendix 29

The Impact of Video games on Ultrasound-Guided Regional Anaesthesia (UGRA) skills.

Participant number [ ]   *Study period* [ ]

Errors - Please tick each time an error is made

Needle advanced without visualisation of needle tip
Failure to identify the target
Failure to recognise orientation of probe with the image screen
Unintentional probe movement
Target malpositioned on screen, including incorrect depth selection
Attention focussed on hand and not image as needle advanced

Total number of errors [ ]   Box A

Number of needle passes [ ]   Box B

Image quality scale (select one)

Scale	Image quality	Description
0	Ideal	• Olive viewed on highest resolution & all contents visualised • Whole needle seen as contacted wall
1	Good	• Olive imaged well enough to define all aspects of contents • Needle tip & part of shaft seen as contacted wall
2	Satisfactory	• Olive visualised so at least wall was identifiable • Needle tip seen as contacted wall
3	poor	• Olive wall was only partly visualised • Needle contact was identifiable only by tissue distortion

0	Ideal
1	Good
2	Satisfactory
3	Poor

[ ]   Box C

Composite Error Score
(Box A + Box B + Box C)   [ ]

Time to perform task [ ]   Mins.

Appendix 30

The Impact of Video games on Ultrasound-Guided Regional Anaesthesia (UGRA) skills.

Participant Number [　　　]   **Study Period** [　　　]

Global Rating Scale for ultrasound-guided regional anaesthesia.

Please tick and score each of the 7 items below appropriately;

	1	2	3	4	5
Preparation for procedure	Did not organise equipment well Has to stop procedure frequently to prepare equipment. [ ]	[ ]	Equipment generally organised. Occasionally has to stop and prepare items. [ ]	[ ]	All equipment neatly organised, prepared and ready for use. [ ]
Time and motion	Many unnecessary moves. [ ]	[ ]	Efficient time/motion but some unnecessary moves. [ ]	[ ]	Clear economy of movement and maximum efficiency. [ ]
Instrument handling	Repeatedly makes tentative or awkward moves with instruments. [ ]	[ ]	Competent use of instruments but occasionally appeared stiff or awkward. [ ]	[ ]	Fluid moves with instruments and no awkwardness. [ ]
Flow of Procedure	Frequently stopped procedure and seemed unsure of next move. [ ]	[ ]	Demonstrated some forward planning with reasonable progression of procedure. [ ]	[ ]	Obviously planned course of procedure with effortless flow from one move to the next. [ ]
Image quality	Inadequate image. Target partially visualised. Needle movement only identifiable by tissue distortion. [ ]	[ ]	Good image. Most of target visualised. Needle tip and part of needle shaft visible. [ ]	[ ]	Outstanding image. Unequivocal image with complete structure visualised. Whole of needle visible. [ ]
Knowledge of procedure	Deficient knowledge. [ ]	[ ]	Knew all important steps of procedure [ ]	[ ]	Demonstrated familarity with all aspects of procedure [ ]
Overall performance	Very poor. [ ]	[ ]	Competent. [ ]	[ ]	Clearly superior. [ ]

CHAPTER – 5

Conclusions

CONCLUSION

Among the various methods for assessment of procedural skills, this thesis attempts to enlighten the most distinctive selection of techniques for assessing trainees, taking into consideration the strengths and weaknesses of each method that is currently being used in anaesthetic practice.

This thesis tries to filter out the methods that are being used in the field of anaesthesia. For example *logbooks or procedural lists* are more appropriate for providing information about any expected prospects within a training programme. As far as *psychometric or aptitude testing* is concerned there seems negligible proof to endorse its usage in anaesthesia. However, another method described as *cusum analysis* has been found to conduct a robust statistical appraisal of procedural expertise. The only downside with its usage is that it relies upon either self-reported performance or repeated direct observations and may require large numbers of performances to establish competency. Modern tools such as *motion analysis systems* may concentrate on manual dexterity throughout the procedural tasks but needs further validation before its implementation in the assessment of procedural skills in anesthesia can be established.

Another interesting method described as *patient simulation*, is going through an evolution phase in anaesthesia training programmes. This method however remains controversial for evaluation, though predicted to be very informative if proved and tested in likely future studies. The technical challenge is to develop and test simulators

that are accurate enough and have an appropriate scope of challenges so that it could be employed for high-stakes assessment.

Objective structured assessment of technical skills (OSATS) may become a significant part of the future of procedural skills assessment in anaesthesia if appropriate part-task simulators can be validated. However, currently there is not satisfactory evidence to endorse that anaesthetic trainees are assessed in procedural skills using simulators, particularly when these skills can be consistently evaluated by direct observation of performance on patients.

According to the literature, the most valid, reliable and comprehensive assessment tool for assessing the procedural skill in anaesthesia so far seems to be the *checklists* and the *global rating scale (GRS)*. There is even much more emphasis on the usage of these tools in combination rather than used alone in the setting of medical educational research and is therefore considered to be the gold standard. This thesis further validates the proficiency of GRS and checklist when merged together to assess the performances in various modes.

The objectives of our *first study* were to identify if visuospatial ability could predict technical performance of an ultrasound – guided needle task by novice operators, and also to describe how emotional state, intelligence and fear of failure could have impact on this. This is important because service efficiencies, patient safety agenda and the European working time regulations have all reduced clinical learning opportunities in anaesthesia, which could compromise patient safety. One solution in regional anaesthesia, would be to permit early, targeted training for anaesthetists who

are predicted to perform poorly. Evidence revealed that visuospatial ability correlates positively with novice performance of simple laparoscopic tasks.

We concluded that mental rotation test (MRT), which is one of the visuospatial assessment predicted novice performance of an ultrasound-guided needling task on a turkey-breast model, and as a trait measure could be used as a tool to focus training resources on less able individuals. In addition this novice performance was adversely affected by anxiety. Therefore both may prove useful in directing targeted training in UGRA. Future studies are desirable to examine whether exclusive training could alter visuospatial ability and thus improve skills acquirement in UGRA. In addition the sensitivity and specificity of MRT, as a screening tool should be assessed.

In the *second study* we aimed to examine the construct validity and reliability of a new assessment tool for the performance of ultrasound – guided regional anaesthesia, by examining whether it can adequately differentiate between performance levels in anaesthetists across the spectrum of expertise. The new tool, which was developed using a modified Delphi method, consisted of a criterion – referenced checklist with global rating scale to assess the performance of ultrasound – guided regional anaesthesia. The authors had established high face validity and concurrent validity of the assessment tool. However, they did not attempt to demonstrate the construct validity and reliability of the assessment tool.

The ability to perform practical procedures safely and efficiently is of paramount importance to the practice of anaesthesia. The concept is important because the last two decades have borne witness to rapid developments in the practice of ultrasound –

guided regional anaesthesia (UGRA). Many distinct national and international regional anaesthesia societies have published a number of educational core curriculums and structured training guidelines to teach and train these new skills in clinical practice. However, the development of complementary assessment tools has lagged behind. Therefore a method of evaluation is needed to ascertain the competence of junior residents in performing the new skills.

Thus, the study concluded that the assessment tool was able to discriminate between different levels of experience and expertise of anaesthetists in the performance of ultrasound – guided regional anaesthesia. Further research will be necessary to assess the effectiveness of this assessment tool in anaesthesia residency programmes.

The ambition of our *third and final study* was to explore whether playing video games will predict psychomotor performance of an ultrasound – guided needling task or could impact on tests of visuospatial ability and also examine the correlation between these innate abilities.

Video games have become an integral part of global culture. In today's society they are regarded as one of the essential genre of leisure activity. Video games are renowned for negative correlations associated with it. However, there are a number of positive benefits of video games. For many years there has been an association between video games and surgical aptitude. There is sufficient evidence available, which suggest that a positive correlation exists between video gaming and generalised surgical ability, including laparoscopic skills. The evidence suggest that gamers have better laparoscopic skills and video games could be used as a cost – effective and fun

way to train basic laparoscopic skills in surgical residents. However, to date within regional anaesthesia none of the studies have evaluated the impact of video games on UGRA skills.

The study concluded that 'gamers' performed significantly better than 'non – gamers', and thus predicted the psychomotor performance of an ultrasound – guided needle task. There is an association of video game experience with UGRA skills. However, the predictive implication of these inferences is uncertain. Future trials should be more consistent and adequately powered, so that stronger evidence on this topic can be accomplished.

FUTURE IMPLICATIONS

Ultrasound-guided regional anaesthesia (UGRA) involves manual dexterity, hand – eye coordination, and sonographic interpretation. Expertise of these skills permits physicians to keep neural targets and the needle in constant view while evading neighbouring structures during block performance. There are three groups of learners: those who have these skills, those who will get it with some practice, and those who will always struggle. What if we could identify those who are destined to struggle before they encounter their first patient and get them prepared with training and simulated practice?[285]

We recognized that although the visuospatial (MRT) assessment may predict novice performance on an ultrasound task but it does not specify at what score, performance could be described as acceptable. It is exciting to reflect about how visuospatial assessments may be used in the future to recognize learners who could benefit from more supervision, additional practice, or even a diverse kind of learning experience. An equivalent assessment could be employed following residency training to help practicing anaesthesiologists recognize their proficiency level, their own areas of strengths and weakness, thus facilitating them to target their own learning efforts.

There is a possibility that the educators will be able to design exclusive workshops that fast-track the speed with which those learners with lesser visuospatial abilities are capable of acquiring better skills.

It is known that mental rotation skills' training is associated with improved performance of mental rotation tasks.[286] Furthermore, it has been shown that mental rotation skills training transfers to improved novice performance of fundamental laparoscopic skill tasks.[287] Our study group has demonstrated that mental rotation ability correlated significantly with technical performance of an ultrasound-guided needle task in undergraduates.[288] However, the causal nature of this relationship is yet to be elucidated.

Within the time and resource constraints of postgraduate medical training, it would be advantageous to optimise expertise acquisition of practical skills with cheap, self-directed educational intervention. Therefore as an implication strategy in a subsequent study our immediate objective is to determine whether mental rotation skills training is associated with improved performance of an ultrasound-guided needle task. Our hypothesis is that improved mental rotation skills will translate to better technical performance of an ultrasound-guided needle task.

The ability to actually measure the influence of stress and anxiety on performance is somewhat new to anaesthesiology and is also very exciting. It would be fascinating to assess learner stress levels as they practice and developed their technical skills and eventually use this evidence to establish whether a degree of comfort can be achieved through improved visuospatial skills or simulated practice.

Training programs in medicine are becoming liable to different regulatory bodies to make sure that physicians in training are objectively assessed for their skill acquirement. The validation of a criterion-referenced assessment tool[219] is a next step towards achieving this goal.

The assessment tool could also be utilized in diagnosing the deficiencies in any training program and in particular of any individual trainee who might need further teaching and training. This may improve the way the trainees are being appraised and recognize phases of training that needed further improvement. Later the same tool can be used to record trainees' progress. The assessment tool can also be utilized to identify gaps in learning and provide objective feedback that will facilitate trainees to improve their skill acquirement. The most important aspect of this assessment tool is that it has the potential to be used for all types of UGRA procedures.

Future study is required to evaluate the efficacy of the assessment checklist and global rating scale as training tools in anaesthesia residency programmes. In addition to that these assessment tools when used in different programs should be further studied for its learning outcomes and it's feasibility for use in the form of feedbacks from students and instructors.

Simulators are effective in laparoscopic skill acquisition, but they are costly, and not easily and readily available. On the contrary video games are economical, could be used for teaching and are often chosen over lectures.[278,279] Due to the broad and continued charm of video games, it seems logical to investigate and describe their positive effects for education, skill acquisition and maintenance.

We have shown that video-gamers demonstrate enhanced mental rotation skills and improved technical performance of a simple ultrasound-guided needle task; this may imply that video game playing can be used as an adjunct to facilitate or accelerate UGRA expertise acquisition. However, it would first be useful to define and quantify what features of video game playing translate to improved UGRA technical performance. In this way, we will be in a better position to develop video-game play, which is specifically designed to enhance UGRA technical performance. As a first step, we would like to confirm the findings of this observational study in a randomized controlled trial of video game play in individuals with low baseline assessments of UGRA task performance (CES, GRS).

REFERENCES

References:

1.   Greher M, Retzl G, Niel P, Kamolz L, Marhofer P, Kapral S: Ultrasonographic assessment of topographic anatomy in volunteers suggests a modification of the infraclavicular vertical brachial plexus block. Br J Anaesth 2002; 88: 632-6

2.   Marhofer P, Schrogendorfer K, Koinig H, Kapral S, Weinstabl C, Mayer N: Ultrasonographic guidance improves sensory block and onset time of three-in-one blocks. Anesth Analg 1997; 85: 854-7

3.   Marhofer P, Sitzwohl C, Greher M, Kapral S: Ultrasound guidance for infraclavicular brachial plexus anaesthesia in children. Anaesthesia 2004; 59: 642-6

4.   Retzl G, Kapral S, Greher M, Mauritz W: Ultrasonographic findings of the axillary part of the brachial plexus. Anesth Analg 2001; 92: 1271-5

5.   Sandhu NS, Capan LM: Ultrasound-guided infraclavicular brachial plexus block. Br J Anaesth 2002; 89: 254-9

6.   Williams SR, Chouinard P, Arcand G, Harris P, Ruel M, Boudreault D, Girard F: Ultrasound guidance speeds execution and improves the quality of supraclavicular block. Anesth Analg 2003; 97: 1518-23

7.   Marhofer P, Greher M, Kapral S: Ultrasound guidance in regional anaesthesia. Br J Anaesth 2005; 94: 7-17

8.   Chan VW, Perlas A, Rawson R, Odukoya O: Ultrasound-guided supraclavicular brachial plexus block. Anesth Analg 2003; 97: 1514-7

9.   Marhofer P, Schrogendorfer K, Andel H, Koinig H, Girsch W, Kapral S, Mayer N: [Combined sciatic nerve-3 in 1 block in high risk patient]. Anasthesiol Intensivmed Notfallmed Schmerzther 1998; 33: 399-401

10.   Grau T, Leipold RW, Conradi R, Martin E, Motsch J: Efficacy of ultrasound imaging in obstetric epidural anesthesia. J Clin Anesth 2002; 14: 169-75

11.   Willschke H, Marhofer P, Bosenberg A, Johnston S, Wanzel O, Cox SG, Sitzwohl C, Kapral S: Ultrasonography for ilioinguinal/iliohypogastric nerve blocks in children. Br J Anaesth 2005; 95: 226-30

12.   Griffin J, Nicholls B: Ultrasound in regional anaesthesia. Anaesthesia; 65 Suppl 1: 1-12

13.   Chan VS, Perlas A, McCartney CL, Brull R, Xu D, Abbas S: Ultrasound guidance improves success rate of axillary brachial plexus block. Canadian Journal of Anesthesia 2007; 54: 176-182

14.   Orebaugh SL, Williams BA, Kentor ML: Ultrasound guidance with nerve stimulation reduces the time necessary for resident peripheral nerve blockade. Reg Anesth Pain Med 2007; 32: 448-54

15.   Perlas A, Brull R, Chan VW, McCartney CJ, Nuica A, Abbas S: Ultrasound guidance improves the success of sciatic nerve block at the popliteal fossa. Reg Anesth Pain Med 2008; 33: 259-65

16.   Casati A, Danelli G, Baciarello M, Corradi M, Leone S, Di Cianni S, Fanelli G: A prospective, randomized comparison between ultrasound and nerve

stimulation guidance for multiple injection axillary brachial plexus block. Anesthesiology 2007; 106: 992-6

17. Sandhu NS, Manne JS, Medabalmi PK, Capan LM: Sonographically guided infraclavicular brachial plexus block in adults: a retrospective analysis of 1146 cases. J Ultrasound Med 2006; 25: 1555-61

18. Marhofer P, Schrogendorfer K, Wallner T, Koinig H, Mayer N, Kapral S: Ultrasonographic guidance reduces the amount of local anesthetic for 3-in-1 blocks. Reg Anesth Pain Med 1998; 23: 584-8

19. Hopkins PM: Ultrasound guidance as a gold standard in regional anaesthesia. Br J Anaesth 2007; 98: 299-301

20. Ramlogan R, Manickam B, Chan VW, Liang L, Adhikary SD, Liguori GA, Hargett MJ, Brull R: Challenges and training tools associated with the practice of ultrasound-guided regional anesthesia: a survey of the American society of regional anesthesia and pain medicine. Reg Anesth Pain Med; 35: 224-6

21. Peterson MK, Millar FA, Sheppard DG: Ultrasound-guided nerve blocks. Br J Anaesth 2002; 88: 621-4

22. Smith HM, Kopp SL, Jacob AK, Torsher LC, Hebl JR: Designing and implementing a comprehensive learner-centered regional anesthesia curriculum. Reg Anesth Pain Med 2009; 34: 88-94

23. Brydges R, Carnahan H, Backstein D, Dubrowski A: Application of motor learning principles to complex surgical tasks: searching for the optimal practice schedule. J Mot Behav 2007; 39: 40-8

24.   Sites BD, Chan VW, Neal JM, Weller R, Grau T, Koscielniak-Nielsen ZJ, Ivani G: The American Society of Regional Anesthesia and Pain Medicine and the European Society Of Regional Anaesthesia and Pain Therapy Joint Committee recommendations for education and training in ultrasound-guided regional anesthesia. Reg Anesth Pain Med 2009; 34: 40-6

25.   Trunkey DD, Botney R: Assessing competency: a tale of two professions. J Am Coll Surg 2001; 192: 385-95

26.   Gaba DM: Structural and organizational issues in patient safety: A comparison of health care to other high-hazard industries. California Management Rev. 2001; 43: 83-102

27.   Kohn K, Corrigan J, Donaldson M: To err is human: building a safer health system. Washington, 1999

28.   Jagsi R, Kitch BT, Weinstein DF, Campbell EG, Hutter M, Weissman JS: Residents report on adverse events and their causes. Arch Intern Med 2005; 165: 2607-13

29.   Stockwell DC, Pollack MM, Turenne WM, Slonim AD: Leadership and management training of pediatric intensivists: how do we gain our skills? Pediatr Crit Care Med 2005; 6: 665-70

30.   Wu AW, Folkman S, McPhee SJ, Lo B: Do house officers learn from their mistakes? JAMA 1991; 265: 2089-94

31.   Weinger MB, Ancoli-Israel S: Sleep deprivation and clinical performance. JAMA 2002; 287: 955-7

32.   Daugherty SR, Baldwin DC, Rowley BD: Learning, satisfaction, and mistreatment during medical internship: a national survey of working conditions. JAMA 1998; 279: 1194-9

33.   Robbins J, Gottlieb F: Sleep deprivation and cognitive testing in internal medicine house staff. West J Med 1990; 152: 82-6

34.   Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, Lilly CM, Stone PH, Lockley SW, Bates DW, Czeisler CA: Effect of reducing interns' work hours on serious medical errors in intensive care units. N Engl J Med 2004; 351: 1838-48

35.   Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, Speizer FE, Czeisler CA: Impact of extended-duration shifts on medical errors, adverse events, and attentional failures. PLoS Med 2006; 3: e487

36.   Reznick RK, MacRae H: Teaching surgical skills--changes in the wind. N Engl J Med 2006; 355: 2664-9

37.   Sexton J, Thomas J, Pronovost P: The context of care and the patient care team: The Safety Attitudes Questionnaire. Building a better delivery system: a new engineering health care partnership. Washington: The National Academy of Sciences., 2005

38.   Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, Roberts PR, Thomas EJ: The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. BMC Health Serv Res 2006; 6: 44

39.   Thomas EJ, Sexton JB, Neilands TB, Frankel A, Helmreich RL: Correction: The effect of executive walk rounds on nurse safety climate attitudes: A randomized trial of clinical units [ISRCTN85147255]. BMC Health Serv Res 2005; 5: 46

40.   Volpp KG, Grande D: Residents' suggestions for reducing errors in teaching hospitals. N Engl J Med 2003; 348: 851-5

41.   Forum H: Patient safety in the ED. Hosp Health Netw 2006;80 http://www.hhnmag.com/hhnmag/jsp/articledisplay.jsp?dcrpath=HHNMAG/Article/data, 2006

42.   Trzeciak S, Rivers EP: Emergency department overcrowding in the United States: an emerging threat to patient safety and public health. Emerg Med J 2003; 20: 402-5

43.   Makary M, Sexton J, Freischlag J, al e: Patient safety in surgery. Ann Surg 2006, 243:628-32; discussion 632-5, 2006

44.   Christian CK, Gustafson ML, Roth EM, Sheridan TB, Gandhi TK, Dwyer K, Zinner MJ, Dierks MM: A prospective study of patient safety in the operating room. Surgery 2006; 139: 159-73

45.   Stockwell DC, Slonim AD: Quality and safety in the intensive care unit. J Intensive Care Med 2006; 21: 199-210

46.   Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, Newhouse JP, Weiler PC, Hiatt HH: Incidence of adverse events and negligence

in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991; 324: 370-6

47.   Reason J: Understanding adverse events: human factors. Qual Health Care, 1995, pp 80-9

48.   Pronovost PJ, Thompson DA, Holzmueller CG, Lubomski LH, Dorman T, Dickman F, Fahey M, Steinwachs DM, Engineer L, Sexton JB, Wu AW, Morlock LL: Toward learning from patient safety reporting systems. J Crit Care 2006; 21: 305-15

49.   McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G, Collins CH: Confidential inquiry into quality of care before admission to intensive care. BMJ 1998; 316: 1853-8

50.   Baldwin PJ, Dodd M, Wrate RM: Junior doctors making mistakes. Lancet 1998; 351: 804

51.   Mason WT, Strike PW: See one, do one, teach one--is this still how it works? A comparison of the medical and nursing professions in the teaching of practical procedures. Med Teach 2003; 25: 664-6

52.   Smith CC, Gordon CE, Feller-Kopman D, Huang GC, Weingart SN, Davis RB, Ernst A, Aronson MD: Creation of an innovative inpatient medical procedure service and a method to evaluate house staff competency. J Gen Intern Med 2004; 19: 510-3

53.   Lighthall G: The IMPES course: toward better outcomes through simulation-based multidisciplinary team training. 1st ed. Des Plaines: Society of Critical Care Medicine., 2004

54.   Cooper JB, Newbower RS, Long CD, McPeek B: Preventable anesthesia mishaps: a study of human factors. Anesthesiology 1978; 49: 399-406

55.   Lighthall GK, Barr J, Howard SK, Gellar E, Sowb Y, Bertacini E, Gaba D: Use of a fully simulated intensive care unit environment for critical event management training for internal medicine residents. Crit Care Med 2003; 31: 2437-43

56.   Cox CE, Carson SS, Ely EW, Govert JA, Garrett JM, Brower RG, Morris DG, Abraham E, Donnabella V, Spevetz A, Hall JB: Effectiveness of medical resident education in mechanical ventilation. Am J Respir Crit Care Med 2003; 167: 32-8

57.   Caplan RA, Posner KL, Cheney FW: Effect of outcome on physician judgments of appropriateness of care. JAMA 1991; 265: 1957-60

58.   Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, Hebert L, Newhouse JP, Weiler PC, Hiatt H: The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med 1991; 324: 377-84

59.   Rothschild JM, Landrigan CP, Cronin JW, Kaushal R, Lockley SW, Burdick E, Stone PH, Lilly CM, Katz JT, Czeisler CA, Bates DW: The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Crit Care Med 2005; 33: 1694-700

60.   Wu A, Holzmueller C, Lubomski L, al. e: Development of the ICU safety reporting system, J Patient Saf, 2005, pp 23-32

61.   Thompson DA, Lubomski L, Holzmueller C, Wu A, Morlock L, Fahey M, Dickman F, Dorman T, Pronovost P: Integrating the intensive care unit safety reporting system with existing incident reporting systems. Jt Comm J Qual Patient Saf 2005; 31: 585-93

62.   Wu AW, Pronovost P, Morlock L: ICU incident reporting systems. J Crit Care 2002; 17: 86-94

63.   Harden RM: Developments in outcome-based education. Med Teach 2002; 24: 117-20

64.   ACGME: (Accreditation Council for Graduate Medical Education) (1999) Outcome project [http://www.acgme.org/Outcome/comp2.asp],

65.   Schwarz MR, Wojtczak A: Global minimum essential requirements in medical education: a road towards competence-oriented medical education. Med Teach 2002; 24: 125-9

66.   GMC: Recommendations on Undergraduate Medical Education (London, GMC). 2002

67.   Simpson JG, Furnace J, Crosby J, Cumming AD, Evans PA, Friedman Ben David M, Harden RM, Lloyd D, McKenzie H, McLachlan JC, McPhate GF, Percy-Robb IW, MacPherson SG: The Scottish doctor--learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. Med Teach 2002; 24: 136-43

68.   Donaldson L: Unfinished Business: Proposals for Reform of the Senior House Officer Grade (London, Department of Health), 2002

69.   Joint Committee on Higher Medical Training: Higher Medical Training: Generic Curriculum, 2003

70.   Rowley D, Pitts D, Galasko C: Report of the Working Party on Evaluation of Surgical Competence. [Available from Professor D. Rowley, Department of Orthopaedic Surgery, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY.], 2002

71.   McGaghie WC, et al.: Competency-based Curriculum Development in Medical Education (Geneva, WHO). 1978

72.   Rethans JJ, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T, Lew S, Page GG, Southgate LH: The relationship between competence and performance: implications for assessing practice performance. Med Educ 2002; 36: 901-9

73.   Klass D: Reevaluation of clinical competency. Am J Phys Med Rehabil 2000; 79: 481-6

74.   Flanagan B, Nestel D, Joseph M: Making patient safety the focus: crisis resource management in the undergraduate curriculum. Med Educ 2004; 38: 56-66

75.   Mayer D, Klamen DL, Gunderson A, Barach P: Designing a patient safety undergraduate medical curriculum: the Telluride Interdisciplinary Roundtable experience. Teach Learn Med 2009; 21: 52-8

76. Wakefield A, Attree M, Braidman I, Carlisle C, Johnson M, Cooke H: Patient safety: do nursing and medical curricula address this theme? Nurse Educ Today 2005; 25: 333-40

77. Smith EL, Cronenwett L, Sherwood G: Current assessments of quality and safety education in nursing. Nurs Outlook 2007; 55: 132-7

78. Epstein RM, Hundert EM: Defining and assessing professional competence. JAMA 2002; 287: 226-35

79. Kogan JR, Holmboe ES, Hauer KE: Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. JAMA 2009; 302: 1316-26

80. Carr S: The Foundation Programme assessment tools: an opportunity to enhance feedback to trainees? Postgrad Med J 2006; 82: 576-9

81. Walton MM, Shaw T, Barnet S, Ross J: Developing a national patient safety education framework for Australia. Qual Saf Health Care 2006; 15: 437-42

82. Miller GE: The assessment of clinical skills/competence/performance. Acad Med 1990; 65: S63-7

83. Rodriguez-Paz JM, Kennedy M, Salas E, Wu AW, Sexton JB, Hunt EA, Pronovost PJ: Beyond "see one, do one, teach one": toward a different training paradigm. Postgrad Med J 2009; 85: 244-9

84. Richardson V: The evolution of reflective teaching and teacher education. In: Clift RT, Houston WR, Pugach MC, eds. Encouraging reflective

practice in education. An analysis of issues and programs. New York, Teachers College Press, 1990, pp 3-19

85.   Schon D: The reflective practitioner: how professionals think in action. London, Ashgate Publishing, 1994

86.   Conant J: The Education of American Teachers. New York, McGraw Hill, 1963

87.   Wolf A: Assessing Assessment. Competence Based Assessment. Philadelphia, Open University Press, 1995

88.   Jessop G: Outcomes. NVQ's and the Emerging Model of Education and Training. London, Falmer, 1991

89.   Hager P, Gonczi A: What is competence? Medical Teacher, 1996, pp 15-18

90.   Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C: Shifting paradigms: from Flexner to competencies. Acad Med 2002; 77: 361-7

91.   Shumway JM, Harden RM: Association for Medical Education in Europe. AMEE Guide No. 25: the assessment of learning outcomes for the competent and reflective physician. Med Teach 2003; 25: 569-84

92.   Crossley J, Humphris G, Jolly B: Assessing health professionals. Med Educ 2002; 36: 800-4

93.   Newman J: Selected Discourses on Liberal Knowledge. Dublin: James Duffy 1852, reprinted 1928.,

94.   Glavin RJ, Maran NJ: Development and use of scoring systems for assessment of clinical competence. Br J Anaesth 2002; 88: 329-30

95.   Objectives of Training in Anaesthesia. Melbourne, Australian and New Zealand College of Anaesthetists, 1992

96.   Newble D, Dauphinee D, Dawson Saunders B, MacDonald M, Mullholand H, Page G, Swanson D, Thomson A, Van der Vlueten C: Guidelines for the development of effective and efficient procedures for the assessment of clinical competence. In: Newble D, Jolly B, Wakeford R, eds The Certification and Recertification of Doctors: Issues in the Assessment of Clinical Competence. Cambridge, Cambridge University Press, 1994

97.   Dunn W, Hamilton D, Harden R: Techniques of identifying competencies needed for doctors., Medical Teacher, 1985, pp 15-25

98.   Nuttall D: The validity of assessments. European Journal of Psychology of Education 1987; 2: 109-118

99.   Appraising Doctors and Dentists in Training. A SCOPME working paper for consultation., 1996, pp 8-27

100.   Paget N, Newble D, Saunders N, Du J: Physician assessment pilot study for the Royal Australasian College of Physicians, Journal of Continuing Education in the Health Professions, 1996, pp 103-111

101.   Jolly B, Grant J: The good assessment guide. Joint Centre for Education in Medicine., 1997, pp 13

102.   McKinley RK, Strand J, Ward L, Gray T, Alun-Jones T, Miller H: Checklists for assessment and certification of clinical procedural skills omit essential competencies: a systematic review. Med Educ 2008; 42: 338-49

103.   Kopta JA: The development of motor skills in orthopaedic education. Clin Orthop Relat Res 1971; 75: 80-5

104.   Hamdorf JM, Hall JC: Acquiring surgical skills. Br J Surg 2000; 87: 28-37

105.   Pangaro L: A new vocabulary and other innovations for improving descriptive in-training evaluations. Acad Med 1999; 74: 1203-7

106.   Naik VN, Perlas A, Chandra DB, Chung DY, Chan VW: An assessment tool for brachial plexus regional anesthesia performance: establishing construct validity and reliability. Reg Anesth Pain Med 2007; 32: 41-5

107.   Tetzlaff JE: Assessment of competency in anesthesiology. Anesthesiology 2007; 106: 812-25

108.   Gardner J: Assessment and learning. London, Sage Publications, 2006

109.   Wragg A, Wade W, Fuller G, Cowan G, Mills P: Assessing the performance of specialist registrars. Clin Med 2003; 3: 131-4

110.   Turnbull J, Gray J, MacFadyen J: Improving in-training evaluation programs. J Gen Intern Med 1998; 13: 317-23

111.   Schmidt R, Bjork R: New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. Psychol Sci, 1992, pp 207-17

112.   Xeroulis GJ, Park J, Moulton CA, Reznick RK, Leblanc V, Dubrowski A: Teaching suturing and knot-tying skills to medical students: a randomized controlled study comparing computer-based video instruction and (concurrent and summary) expert feedback. Surgery 2007; 141: 442-9

113.   Hunt RJ: Percent agreement, Pearson's correlation, and kappa as measures of inter-examiner reliability. J Dent Res 1986; 65: 128-30

114.   Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, Fried GM: A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg 2005; 190: 107-13

115.   Muller R, Buttner P: A crtical discussion of intraclass correlation coefficients. Stat Med, 1994, pp 2465-76

116.   Landis J, Koch G: An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977; 33: 363-74

117.   Friedman Z, Katznelson R, Devito I, Siddiqui M, Chan V: Objective assessment of manual skills and proficiency in performing epidural anesthesia--video-assisted validation. Reg Anesth Pain Med 2006; 31: 304-10

118. Thompson WG, Lipkin M, Gilbert DA, Guzzo RA, Roberson L: Evaluating evaluation: assessment of the American Board of Internal Medicine Resident Evaluation Form. J Gen Intern Med 1990; 5: 214-7

119. Holmboe ES, Hawkins RE: Methods for evaluating the clinical competence of residents in internal medicine: a review. Ann Intern Med 1998; 129: 42-8

120. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M: Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 1997; 84: 273-8

121. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W: Testing technical skill via an innovative "bench station" examination. Am J Surg 1997; 173: 226-30

122. Naik VN, Matsumoto ED, Houston PL, Hamstra SJ, Yeung RY, Mallon JS, Martire TM: Fiberoptic orotracheal intubation on anesthetized patients: do manipulation skills learned on a simple model transfer into the operating room? Anesthesiology 2001; 95: 343-8

123. Goff BA, Lentz GM, Lee D, Houmard B, Mandel LS: Development of an objective structured assessment of technical skills for obstetric and gynecology residents. Obstet Gynecol 2000; 96: 146-50

124. Morgan PJ, Cleave-Hogg D, Guest CB: A comparison of global ratings and checklist scores from an undergraduate assessment using an anesthesia simulator. Acad Med 2001; 76: 1053-5

125.   Regehr G, MacRae H, Reznick RK, Szalay D: Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. Acad Med 1998; 73: 993-7

126.   Aggarwal R, Grantcharov T, Moorthy K, Milland T, Darzi A: Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. Ann Surg 2008; 247: 372-9

127.   Sites BD, Spence BC, Gallagher JD, Wiley CW, Bertrand ML, Blike GT: Characterizing novice behavior associated with learning ultrasound-guided peripheral regional anesthesia. Reg Anesth Pain Med 2007; 32: 107-15

128.   Vadodaria BS, Gandhi SD, McIndoe AK: Comparison of four different emergency airway access equipment sets on a human patient simulator. Anaesthesia 2004; 59: 73-9

129.   Ziv A, Wolpe PR, Small SD, Glick S: Simulation-based medical education: an ethical imperative. Acad Med 2003; 78: 783-8

130.   Sulaiman L, Tighe SQ, Nelson RA: Surgical vs wire-guided cricothyroidotomy: a randomised crossover study of cuffed and uncuffed tracheal tube insertion. Anaesthesia 2006; 61: 565-70

131.   Crawford SW, Colt HG: Virtual reality and written assessments are of potential value to determine knowledge and skill in flexible bronchoscopy. Respiration 2004; 71: 269-75

132.   Kestin IG: A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. Br J Anaesth 1995; 75: 805-9

133.   Young A, Miller JP, Azarow K: Establishing learning curves for surgical residents using Cumulative Summation (CUSUM) Analysis. Curr Surg 2005; 62: 330-4

134.   Naik VN, Devito I, Halpern SH: Cusum analysis is a useful tool to assess resident proficiency at insertion of labour epidurals. Can J Anaesth 2003; 50: 694-8

135.   de Oliveira Filho G: The construction of learning curves for basic skills in anaesthetic procedures: an application for the cumulative sum method. Anesth Analg, 2002, pp 411-6

136.   Konrad C: Learning manual skills in anesthesiology: is there a recommended number of cases for anesthetic procedures? Anesth Analg, 1998, pp 635-9

137.   Bould MD, Crabtree NA: Are logbooks of training in anaesthesia a valuable exercise? Br J Hosp Med (Lond) 2008; 69: 236

138.   Wanzel KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD: Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet 2002; 359: 230-1

139.   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: 1468-71

140.   Dashfield AK, Coghill JC, Langton JA: Correlating obstetric epidural anaesthesia performance and psychomotor aptitude. Anaesthesia 2000; 55: 744-9

141.   Dashfield AK, Smith JE: Correlating fibreoptic nasotracheal endoscopy performance and psychomotor aptitude. Br J Anaesth 1998; 81: 687-91

142.   Ault G, Reznick R, MacRae H, Leadbetter W, DaRosa D, Joehl R, Peters J, Regehr G: Exporting a technical skills evaluation technology to other sites. Am J Surg 2001; 182: 254-6

143.   Goff BA, Nielsen PE, Lentz GM, Chow GE, Chalmers RW, Fenner D, Mandel LS: Surgical skills assessment: a blinded examination of obstetrics and gynecology residents. Am J Obstet Gynecol 2002; 186: 613-7

144.   Datta V, Chang A, Mackay S, Darzi A: The relationship between motion analysis and surgical technical assessments. Am J Surg 2002; 184: 70-3

145.   Datta V, Mackay S, Mandalia M, Darzi A: The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. J Am Coll Surg 2001; 193: 479-85

146.   Chin KJ, Tse C, Chan V, Tan JS, Lupu CM, Hayter M: Hand motion analysis using the imperial college surgical assessment device: validation of a novel and objective performance measure in ultrasound-guided peripheral nerve blockade. Reg Anesth Pain Med 2011; 36: 213-9

147.   Smith HM, Kopp SL, Johnson RL, Long TR, Cerhan JH, Hebl JR: Looking into learning: visuospatial and psychomotor predictors of ultrasound-guided procedural performance. Reg Anesth Pain Med 2012; 37: 441-7

148.   National Institute of Clinical Excellence: Ultrasound Guided Regional Nerve Block.   http://www.nice.org.uk/nicemedia/live/11931/42876/42876.pdf (Accessed 20/12/2013). 2009

149.   Association of Anaesthetists of Great Britain and Ireland. Ultrasound in Anaesthesia and Intensive Care: a Guide to Training. London:AAGBI, 2011. http://www.aagbi.org/sites/default/files/Ultrasound%20in%20Anaesthesia%20and%20Intensive%20Care%20-%20A%20Guide%20to%20Training.pdf   (accessed 20/12/2013). 2011

150.   Dalal PG, Dalal GB, Pott L, Bezinover D, Prozesky J, Bosseau Murray W: Learning curves of novice anesthesiology residents performing simulated fibreoptic upper airway endoscopy. Can J Anaesth 2011; 58: 802-9

151.   Watts J FW: Assessment of technical skills., Assessing Clinical Competence. Edited by Nuefeld VR NG. New York, NY, Springer, 1985, pp 259-74

152.   Boyle AM, Santelli JC: Assessing psychomotor skills: the role of the Crawford Small Parts Dexterity Test as a screening instrument. J Dent Educ 1986; 50: 176-9

153.   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 2008; 78: 291-6

154.   Enochsson L, Isaksson B, Tour R, Kjellin A, Hedman L, Wredmark T, Tsai-Fellander L: Visuospatial skills and computer game experience influence the performance of virtual endoscopy. J Gastrointest Surg 2004; 8: 876-82; discussion 882

155.   Enochsson L, Westman B, Ritter EM, Hedman L, Kjellin A, Wredmark T, Fellander-Tsai L: Objective assessment of visuospatial and psychomotor ability and flow of residents and senior endoscopists in simulated gastroscopy. Surg Endosc 2006; 20: 895-9

156.   Francis NK, Hanna GB, Cresswell AB, Carter FJ, Cuschieri A: The performance of master surgeons on standard aptitude testing. Am J Surg 2001; 182: 30-3

157.   Fried MP, Uribe JI, Sadoughi B: The role of virtual reality in surgical training in otorhinolaryngology. Curr Opin Otolaryngol Head Neck Surg 2007; 15: 163-9

158.   Weiner. S: Military Flight Aptitude Tests., 6th Edition. New York, NY: ARCO, 2004

159.   Westman B, Ritter EM, Kjellin A, Torkvist L, Wredmark T, Fellander-Tsai L, Enochsson L: Visuospatial abilities correlate with performance of senior endoscopy specialist in simulated colonoscopy. J Gastrointest Surg 2006; 10: 593-9

160.   Wanzel KR, Hamstra SJ, Caminiti MF, Anastakis DJ, Grober ED, Reznick RK: Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. Surgery 2003; 134: 750-7

161.   Stefanidis D, Korndorffer JR, Jr., Black FW, Dunne JB, Sierra R, Touchard CL, Rice DA, Markert RJ, Kastl PR, Scott DJ: Psychomotor testing predicts rate of skill acquisition for proficiency-based laparoscopic skills training. Surgery 2006; 140: 252-62

162.   Schlickum M, Hedman L, Enochsson L, Henningsohn L, Kjellin A, Fellander-Tsai L: Surgical simulation tasks challenge visual working memory and visual-spatial ability differently. World J Surg 2011; 35: 710-5

163.   Rosenthal R, Hamel C, Oertli D, Demartines N, Gantert WA: Performance on a virtual reality angled laparoscope task correlates with spatial ability of trainees. Indian J Surg 2010; 72: 327-30

164.   Luursema JM, Buzink SN, Verwey WB, Jakimowicz JJ: Visuo-spatial ability in colonoscopy simulator training. Adv Health Sci Educ Theory Pract 2010; 15: 685-94

165.   Jungmann F, Gockel I, Hecht H, Kuhr K, Rasanen J, Sihvo E, Lang H: Impact of perceptual ability and mental imagery training on simulated laparoscopic knot-tying in surgical novices using a Nissen fundoplication model. Scand J Surg 2011; 100: 78-85

166.   Hedman L, Strom P, Andersson P, Kjellin A, Wredmark T, Fellander-Tsai L: High-level visual-spatial ability for novices correlates with performance in a visual-spatial complex surgical simulator task. Surg Endosc 2006; 20: 1275-80

167.   Hassan I, Gerdes B, Koller M, Dick B, Hellwig D, Rothmund M, Zielke A: Spatial perception predicts laparoscopic skills on virtual reality laparoscopy simulator. Childs Nerv Syst 2007; 23: 685-9

168.   Ahlborg L, Hedman L, Murkes D, Westman B, Kjellin A, Fellander-Tsai L, Enochsson L: Visuospatial ability correlates with performance in simulated gynecological laparoscopy. Eur J Obstet Gynecol Reprod Biol 2011; 157: 73-7

169. Eysenck MW: Anxiety and cognitive-task performance. Personality and Individual Differences 1985; 6: 579-586

170. Hardy L, Parfitt G: A catastrophe model of anxiety and performance. Br J Psychol 1991; 82 (Pt 2): 163-78

171. Parfitt G, Hardy L: The effects of competitive anxiety on memory span and rebound shooting tasks in basketball players. J Sports Sci 1993; 11: 517-24

172. Pijpers JR, Oudejans RR, Bakker FC: Anxiety-induced changes in movement behaviour during the execution of a complex whole-body task. Q J Exp Psychol A 2005; 58: 421-45

173. Hardy L: Testing the predictions of the cusp catastrophe model of anxiety and performance. The Sport Psychologist 1996; 10: 140-156

174. Meijer J: Learning potential and anxious tendency: Test anxiety as a bias factor in educational testing. Anxiety, Stress, & Coping 2001; 14: 337-362

175. Schouwenburg HC: Test anxiety unveiled. Test Anxiety: The State of the Art. Moshe Zeidner. Plenum, New York, 1998. 440+xxi pages. European Journal of Personality 1999; 13: 327-329

176. Sarason IG: Stress, anxiety, and cognitive interference: reactions to tests. J Pers Soc Psychol 1984; 46: 929-38

177. Elliot A, Pekrun R: Emotion in the hierarchical model of approach-avoidance achievement motivation. In Schutz, Paul A. (Ed); Pekrun, Reinhard (Ed), (2007). Emotion in education. Educational psychology series., (pp. 57-73). . San Diego, CA, Elsevier Academic Press, xiv, 348 pp., 2007

178. Putwain DW, Symes W: Achievement goals as mediators of the relationship between competence beliefs and test anxiety. Br J Educ Psychol 2012; 82: 207-24

179. Bertua C, Anderson N, Salgado JF: The predictive validity of cognitive ability tests: A UK meta-analysis. Journal of Occupational and Organizational Psychology 2005; 78: 387-409

180. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C: A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. Brain Cogn 1995; 28: 39-58

181. Peters M, Chisholm P, Laeng B: Spatial Ability, Student Gender, and Academic Performance. Journal of Engineering Education 1995; 84: 69-73

182. Voyer D, Voyer S, Bryden MP: Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. Psychol Bull 1995; 117: 250-70

183. Shepard S, Metzler D: Mental rotation: effects of dimensionality of objects and type of task. J Exp Psychol Hum Percept Perform 1988; 14: 3-11

184. Witkin HA, Oltman, P. K., Raskin, E., & Karp, S. (1971). A manual for the embedded figures test. Palo Alto, CA: Consulting Psychologists Press.

185. Witkin HA, Oltman, P. K., Raskin, E., & Karp, S. (1971). Group embedded figures test manual. Palo Alto, CA: Consulting Psychology Press.

186. Heim AW: AH4 Group Test of General Intelligence Manual. Windsor, United Kingdom., NFER-Nelson Publishing Company Limited, 1970

187.   Matthews G, Jones DM, Chamberlain AG: Refining the measurement of mood: The UWIST Mood Adjective Checklist. British Journal of Psychology 1990; 81: 17-42

188.   Matthews G, Joyner L, Gilliland K, Campbell SE, Falconer S, Huggins J: Validation of a comprehensive stress state questionnaire: Towards a state big three. Personality psychology in Europe 1999; 7: 335-350

189.   Caraway K, Tucker CM, Reinke WM, Hall C: Self-efficacy, goal orientation, and fear of failure as predictors of school engagement in high school students. Psychology in the Schools 2003; 40: 417-427

190.   Chamorro-Premuzic T: The Numerical Reasoning Test 20-Items (NRT-20). Goldsmiths: University of London., 2008

191.   Ultrasound-guided needle advancement for novices. https://www.youtube.com/watch?v=5VaXrqjRlVs (Accessed on 10/08/2014). 2013

192.   Beese RC, Lowe S: The use of turkey breast and stuffed olives as a soft tissue model for the teaching and practice of ultrasound guided interventional procedures. European Journal of Ultrasound 1998; 7: 12-12

193.   Sites BD, Gallagher JD, Cravero J, Lundberg J, Blike G: The learning curve associated with a simulated ultrasound-guided interventional task by inexperienced anesthesia residents. Reg Anesth Pain Med 2004; 29: 544-8

194.   Davies T, Townsley P, Jlala H, Dowling M, Bedforth N, Hardman JG, McCahon RA: Novice performance of ultrasound-guided needle advancement:

standard 38-mm transducer vs 25-mm hockey stick transducer. Anaesthesia 2012; 67: 855-61

195.   Hopkins WG: A new view of statistics. Internet Society for Sport Science: http://www.sportsci.org/resource/stats/. 2000

196.   Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420-8

197.   Lexell JE, Downham DY: How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005; 84: 719-23

198.   Ahlborg L, Hedman L, Rasmussen C, Fellander-Tsai L, Enochsson L: Non-technical factors influence laparoscopic simulator performance among OBGYN residents. Gynecological Surgery 2012; 9: 415-420

199.   Linn MC, Petersen AC: Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Dev 1985; 56: 1479-98

200.   Kolozsvari NO, Andalib A, Kaneva P, Cao J, Vassiliou MC, Fried GM, Feldman LS: Sex is not everything: the role of gender in early performance of a fundamental laparoscopic skill. Surg Endosc 2011; 25: 1037-42

201.   Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J: Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. Surg Endosc 2003; 17: 1082-5

202.   Ferguson E, Semper H, Yates J, Fitzgerald JE, Skatova A, James D: The 'dark side' and 'bright side' of personality: when too much conscientiousness and

too little anxiety are detrimental with respect to the acquisition of medical knowledge and skill. PLoS One; 9: e88606

203.   Bridges M, Diamond DL: The financial impact of teaching surgical residents in the operating room. Am J Surg 1999; 177: 28-32

204.   Darzi A, Smith S, Taffinder N: Assessing operative skill. Needs to become more objective. BMJ 1999; 318: 887-8

205.   Cuschieri A, Francis N, Crosby J, Hanna GB: What do master surgeons think of surgical competence and revalidation? Am J Surg 2001; 182: 110-6

206.   Reznick RK: Teaching and testing technical skills. Am J Surg 1993; 165: 358-61

207.   Kopta JA: An approach to the evaluation of operative skills. Surgery 1971; 70: 297-303

208.   Neal JM, Brull R, Chan VW, Grant SA, Horn JL, Liu SS, McCartney CJ, Narouze SN, Perlas A, Salinas FV, Sites BD, Tsui BC: The ASRA evidence-based medicine assessment of ultrasound-guided regional anesthesia and pain medicine: Executive summary. Reg Anesth Pain Med 2010; 35: S1-9

209.   Fingerman M, Benonis JG, Martin G: A practical guide to commonly performed ultrasound-guided peripheral-nerve blocks. Curr Opin Anaesthesiol 2009; 22: 600-7

210.   Regional Anesthesiology and Acute Pain Medicine Fellowship Directors Group. Guidelines for fellowship training in Regional Anesthesiology and Acute Pain Medicine: Second Edition, 2010. Reg Anesth Pain Med 2011; 36: 282-8

211. Tan JS, Chin KJ, Chan VW: Developing a training program for peripheral nerve blockade: the "nuts and bolts". Int Anesthesiol Clin 2010; 48: 1-11

212. Hance J, Aggarwal R, Stanbridge R, Blauth C, Munz Y, Darzi A, Pepper J: Objective assessment of technical skills in cardiac surgery. Eur J Cardiothorac Surg 2005; 28: 157-62

213. Beard J: Workplace-based assessment: the need for continued evaluation and refinement. Surgeon. 2011; 9 Suppl 1:S12-3.

214. Academy of Medical Royal Colleges. Improving Assessment. London, UK: AoMRC Press. 2009

215. Weller JM, Jolly B, Misur MP, Merry AF, Jones A, Crossley JG, Pedersen K, Smith K: Mini-clinical evaluation exercise in anaesthesia training. Br J Anaesth 2009; 102: 633-41

216. Weller JM, Jones A, Merry AF, Jolly B, Saunders D: Investigation of trainee and specialist reactions to the mini-Clinical Evaluation Exercise in anaesthesia: implications for implementation. Br J Anaesth 2009; 103: 524-30

217. Moorthy K, Munz Y, Sarker SK, Darzi A: Objective assessment of technical skills in surgery. BMJ 2003; 327: 1032-7

218. Ben-Menachem E, Ezri T, Ziv A, Sidi A, Brill S, Berkenstadt H: Objective Structured Clinical Examination-based assessment of regional anesthesia skills: the Israeli National Board Examination in Anesthesiology experience. Anesth Analg 2011; 112: 242-5

219. Cheung JJ, Chen EW, Darani R, McCartney CJ, Dubrowski A, Awad IT: The creation of an objective assessment tool for ultrasound-guided regional anesthesia using the Delphi method. Reg Anesth Pain Med 2012; 37: 329-33

220. Hsu C-C, Sandford BA: The Delphi Technique: Making Sense Of Consensus. Pract Assess Res Eval. 2007; 12: 1-8

221. Morgan PJ, Lam-McCulloch J, Herold-McIlroy J, Tarshis J: Simulation performance checklist generation using the Delphi technique. Can J Anaesth 2007; 54: 992-7

222. Clayton MJ: Delphi: a technique to harness expert opinion for critical decision-making tasks in education. Educational Psychology 1997; 17: 373-386

223. Rohan D, Ahern S, Walsh K: Defining an anaesthetic curriculum for medical undergraduates. A Delphi study. Med Teach 2009; 31: 1-5

224. Joynt GM, Loo S, Taylor BL, Margalit G, Christian MD, Sandrock C, Danis M, Leoniv Y, Sprung CL: Chapter 3. Coordination and collaboration with interface units. Recommendations and standard operating procedures for intensive care unit and hospital preparations for an influenza epidemic or mass disaster. Intensive Care Med 2010; 36 (Suppl 1): S21-S31

225. Green B, Jones M, Hughes D, Williams A: Applying the Delphi technique in a study of GPs' information requirements. Health Soc Care Community 1999; 7: 198-205

226. Kopacz DJ, Neal JM, Pollock JE: The regional anesthesia "learning curve". What is the minimum number of epidural and spinal blocks to reach consistency? Reg Anesth 1996; 21: 182-90

227. Konrad C, Schupfer G, Wietlisbach M, Gerber H: Learning manual skills in anesthesiology: Is there a recommended number of cases for anesthetic procedures? Anesth Analg 1998; 86: 635-9

228. Rosenblatt MA, Fishkind D: Proficiency in interscalene anesthesia-how many blocks are necessary? J Clin Anesth 2003; 15: 285-8

229. Norris TE, Cullison SW, Fihn SD: Teaching procedural skills. J Gen Intern Med 1997; 12: S64-70

230. de Oliveira Filho GR: The construction of learning curves for basic skills in anesthetic procedures: an application for the cumulative sum method. Anesth Analg 2002; 95: 411-6

231. Sites BD, Chan VW, Neal JM, Weller R, Grau T, Koscielniak-Nielsen ZJ, Ivani G: The American Society of Regional Anesthesia and Pain Medicine and the European Society of Regional Anaesthesia and Pain Therapy joint committee recommendations for education and training in ultrasound-guided regional anesthesia. Reg Anesth Pain Med 2010; 35: S74-S80.

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

233. Gibbons RD, Baker RJ, Skinner DB: Field articulation testing: a predictor of technical skills in surgical residents. J Surg Res 1986; 41: 53-7

234. Kennedy AM, Boyle EM, Traynor O, Walsh T, Hill AD: Video gaming enhances psychomotor skills but not visuospatial and perceptual abilities in surgical trainees. J Surg Educ 2011; 68: 414-20

235. Hegarty M, Keehner M, Cohen C, Montello DR, Lippa Y: The role of spatial cognition in medicine: Applications for selecting and training professionals. In G. Allen (Ed.), Applied spatial cognition (pp. 285–315). Mahwah, NJ: Lawrence Erlbaum Associates., 2007

236. McVicar J, Niazi AU, Murgatroyd H, Chin KJ, Chan VW: Novice performance of ultrasound-guided needling skills: effect of a needle guidance system. Reg Anesth Pain Med 2015; 40: 150-3

237. Mercaldi CJ, Lanes SF: Ultrasound guidance decreases complications and improves the cost of care among patients undergoing thoracentesis and paracentesis. Chest 2013; 143: 532-8

238. Romero R, Jeanty P, Reece EA, Grannum P, Bracken M, Berkowitz R, Hobbins JC: Sonographically monitored amniocentesis to decrease intraoperative complications. Obstet Gynecol 1985; 65: 426-30

239. Rasmussen SN, Holm HH, Kristensen JK, Barlebo H: Ultrasonically-guided liver biopsy. Br Med J 1972; 2: 500-2

240. Goldfinger M, Rothberg R, Stoll S: Sonographic guidance of thyroid needle biopsy. Can Assoc Radiol J 1986; 37: 186-8

241. Slater RJ, Castanelli DJ, Barrington MJ: Learning and teaching motor skills in regional anesthesia: a different perspective. Reg Anesth Pain Med 2014; 39: 230-9

242. Ritter EM, McClusky DA, 3rd, Gallagher AG, Enochsson L, Smith CD: Perceptual, visuospatial, and psychomotor abilities correlate with duration of training required on a virtual-reality flexible endoscopy simulator. Am J Surg 2006; 192: 379-84

243. T. H: Video game sales blast toward record this holiday season. USA Today. December 23, 2002. http://usatoday30.usatoday.com/tech/techreviews/games/2002-12-23-game-sales_x.htm. Accessed 26 April 2015.

244. Gentile DA, Lynch PJ, Linder JR, Walsh DA: The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. J Adolesc 2004; 27: 5-22

245. Strauss RS, Knight J: Influence of the home environment on the development of obesity in children. Pediatrics 1999; 103: e85

246. Stettler N, Signer TM, Suter PM: Electronic games and environmental factors associated with childhood obesity in Switzerland. Obes Res 2004; 12: 896-903

247. Brasington R: Nintendinitis. N Engl J Med 1990; 322: 1473-4

248. R. L: Nintendo issues game gloves. http://www.gamespot.com/news/2541755.html. Accessed February 24, 2015.

249.   Kasteleijn-Nolst Trenite DG, da Silva AM, Ricci S, Binnie CD, Rubboli G, Tassinari CA, Segers JP: Video-game epilepsy: a European study. Epilepsia 1999; 4: 70-4

250.   Lynch P: Type-A Behavior, Hostility, and Cardiovascular Function at Rest and after Playing Video Games in Teenagers. Psychosomatic Medicine 1994; 56: 152-152

251.   Lynch P: Type-A Behavior and stress hormones at rest and after playing violent video games in teenagers. Psychosomatic Medicine 1999; 61: 113

252.   Green CS, Bavelier D: Effect of action video games on the spatial distribution of visuospatial attention. J Exp Psychol Hum Percept Perform 2006; 32: 1465-78

253.   Green CS, Bavelier D: Action-video-game experience alters the spatial resolution of vision. Psychol Sci 2007; 18: 88-94

254.   Yuji H: Computer games and information-processing skills. Percept Mot Skills 1996; 83: 643-7

255.   Griffith JL, Voloschin P, Gibb GD, Bailey JR: Differences in eye-hand motor coordination of video-game users and non-users. Percept Mot Skills 1983; 57: 155-8

256.   Nielsen U, Dahl R, White RF, Grandjean P: [Computer assisted neuropsychological testing of children]. Ugeskr Laeger 1998; 160: 3557-61

257.   De Lisi R, Wolford JL: Improving children's mental rotation accuracy with computer game playing. J Genet Psychol 2002; 163: 272-82

258.   Green CS, Bavelier D: Action video game modifies visual selective attention. Nature 2003; 423: 534-7

259.   Rosser JC, Jr., Lynch PJ, Cuddihy L, Gentile DA, Klonsky J, Merrell R: The impact of video games on training surgeons in the 21st century. Arch Surg 2007; 142: 181-6

260.   Tsai CL, Heinrichs WL: Acquisition of Eye-hand Coordination Skills for Videoendoscopic Surgery. J Am Assoc Gynecol Laparosc 1994; 1 (4, Part 2): S37.

261.   Van Hove C, Perry KA, Spight DH, Wheeler-Mcinvaille K, Diggs BS, Sheppard BC, Jobe BA, O'Rourke RW: Predictors of technical skill acquisition among resident trainees in a laparoscopic skills education program. World J Surg 2008; 32: 1917-21

262.   Shane MD, Pettitt BJ, Morgenthal CB, Smith CD: Should surgical novices trade their retractors for joysticks? Videogame experience decreases the time needed to acquire surgical skills. Surg Endosc 2008; 22: 1294-7

263.   Miskry T, Magos T, Magos A: If you're no good at computer games, don't operate endoscopically! Gynaecological Endoscopy 2002; 11: 345-347

264.   Li R, Polat U, Makous W, Bavelier D: Enhancing the contrast sensitivity function through action video game training. Nat Neurosci 2009; 12: 549-51

265.   Intelligence D: Global video game software industry to reach $100B in 2019.   http://www.dfcint.com/wp/dfc-intelligence-forecasts-global-video-game-software-industry-to-reach-100b-in-2019/. (Accessed on 05th May 2015), 2015

266. Appleget JA, Illingworth JE: Land Warrior Training Initiative. http://www.simulationaustralasia.com/mysimaust/show_abstract/833. (Accessed on 05th May 2015), 2001

267. Walker T: Ubi Soft licenses its tactical action game to the US military to help train soldiers. http://www.gamespot.com/articles/us-department-of-defense-to-use-rogue-spear/1100-2809207/. (Accessd on 05th May 2015). 2001

268. Ou Y, McGlone ER, Camm CF, Khan OA: Does playing video games improve laparoscopic skills? Int J Surg 2013; 11: 365-9

269. Sharma D, Shaban A, Riddell A, Kalsi V, Arya M, Grange P: Video-games station or minimally invasive skills training station? BJU Int 2009; 104: 159-60

270. Lynch J, Aughwane P, Hammond TM: Video games and surgical ability: a literature review. J Surg Educ 2010; 67: 184-9

271. Rosenberg BH, Landsittel D, Averch TD: Can video games be used to predict or improve laparoscopic skills? J Endourol 2005; 19: 372-6

272. Badurdeen S, Abdul-Samad O, Story G, Wilson C, Down S, Harris A: Nintendo Wii video-gaming ability predicts laparoscopic skill. Surg Endosc 2010; 24: 1824-8

273. Schlickum MK, Hedman L, Enochsson L, Kjellin A, Fellander-Tsai L: Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. World J Surg 2009; 33: 2360-7

274.   Hogle NJ, Widmann WD, Ude AO, Hardy MA, Fowler DL: Does training novices to criteria and does rapid acquisition of skills on laparoscopic simulators have predictive validity or are we just playing video games? J Surg Educ 2008; 65: 431-5

275.   Madan AK, Frantzides CT, Park WC, Tebbit CL, Kumari NV, O'Leary PJ: Predicting baseline laparoscopic surgery skills. Surg Endosc 2005; 19: 101-4

276.   Nomura T, Miyashita M, Shrestha S, Makino H, Nakamura Y, Aso R, Yoshimura A, Shimura T, Akira S, Tajiri T: Can interview prior to laparoscopic simulator training predict a trainee's skills? J Surg Educ 2008; 65: 335-9

277.   van Dongen KW, Verleisdonk EJ, Schijven MP, Broeders IA: Will the Playstation generation become better endoscopic surgeons? Surg Endosc 2011; 25: 2275-80

278.   Mann BD, Eidelson BM, Fukuchi SG, Nissman SA, Robertson S, Jardines L: The development of an interactive game-based tool for learning surgical management algorithms via computer. Am J Surg 2002; 183: 305-8

279.   Meterissian S, Liberman M, McLeod P: Games as teaching tools in a surgical residency. Med Teach 2007; 29: e258-60

280.   Verdaasdonk EG, Dankelman J, Schijven MP, Lange JF, Wentink M, Stassen LP: Serious gaming and voluntary laparoscopic skills training: a multicenter study. Minim Invasive Ther Allied Technol 2009; 18: 232-8

281.   Jalink MB, Goris J, Heineman E, Pierie JP, ten Cate Hoedemaker HO: The effects of video games on laparoscopic simulator skills. Am J Surg 2014; 208: 151-6

282.   Bokhari R, Bollman-McGregor J, Kahoi K, Smith M, Feinstein A, Ferrara J: Design, development, and validation of a take-home simulator for fundamental laparoscopic skills: using Nintendo Wii for surgical training. Am Surg 2010; 76: 583-6

283.   Curet M: Commentary on "The impact of video games on training surgeons in the 21st century". Arch Surg 2007; 142: 186

284.   Madan AK, Harper JL, Frantzides CT, Tichansky DS: Nonsurgical skills do not predict baseline scores in inanimate box or virtual-reality trainers. Surg Endosc 2008; 22: 1686-9

285.   Kopp SL, Rathmell JP: Acquiring New Technical Skills and Aptitude for Mental Rotation. Anesthesiology 2015; 123: 993-4

286.   Kail R, Park YS: Impact of practice on speed of mental rotation. J Exp Child Psychol 1990; 49: 227-44

287.   Stransky D, Wilcox LM, Dubrowski A: Mental rotation: cross-task training and generalization. J Exp Psychol Appl 2010; 16: 349-60

288.   Shafqat A, Ferguson E, Thanawala V, Bedforth NM, Hardman JG, McCahon RA: Visuospatial Ability as a Predictor of Novice Performance in Ultrasound-guided Regional Anesthesia. Anesthesiology 2015; 123: 1188-97