A PROFICIENCY-BASED TECHNICAL SKILLS CURRICULUM FOR LAPAROSCOPIC SURGERY

Rajesh Aggarwal

Imperial College London

A thesis submitted for the Degree of
Doctor of Philosophy at the
University of London
April 2008

Dedicated to my parents

for their sacrifices and indeterminate support

Table of Contents

List of Tables and Figures

ACKNOWLEDGEMENTS

I am greatly indebted to Professor Sir Ara Darzi, my supervisor, for providing me with the opportunity to undertake this period of research, and for his guidance, support and encouragement.

I would also like to thank the following without whom this piece of work would not have been possible: Thanos Athanasiou, Indran Balasundaram, Fernando Bello, Camilo Boza, Aristotelis Dosis, Teodor Grantcharov, Julian Hance, Julian Leong, Thor Milland, Krishna Moorthy, Natalia Ognjenovic, Parv Sains and Jonnie Ward.

Finally, I would like to thank the numerous surgeons, medical students and support staff that willingly helped me to perform my studies.

DECLARATION

I declare that the experiments undertaken in this thesis were performed by myself under the supervision of Professor Sir Ara Darzi. This thesis has been composed by myself and is a record of work which has not been previously submitted for a higher degree.

Rajesh Aggarwal

November 2007

STATEMENT

This is to certify that Rajesh Aggarwal has completed nine terms of experimental research, and that he is qualified to submit the following thesis for the Degree of Doctor of Philosophy.

Professor Sir Ara Darzi

(Research supervisor)

November 2007

Papers Originated from the Thesis

1. **Aggarwal R**, Moorthy K, Darzi A. Laparoscopic skills training and assessment. ***Br J Surg*** 2004; 91(12):1549-1558.
2. **Aggarwal R**, Grantcharov TP, Darzi A. A framework for systematic training and assessment of technical skills – STATS. ***J Am Coll Surg*** 2007; 204(4): 697-705.
3. **Aggarwal R**, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. ***Am J Surg*** 2006; 191(1): 128-133.
4. **Aggarwal R**, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen V, Funch-Jensen P, Darzi A. An evidence-based virtual reality training program for laparoscopic surgery. ***Ann Surg*** 2006; 244(2): 310-4.
5. **Aggarwal R**, Grantcharov T, Moorthy K, Milland T, Papasavas P, Dosis A, Bello F, Darzi A. An evaluation of the feasibility, validity and reliability of laparoscopic skills assessment in the operating room. ***Ann Surg*,** 2007; 245(6): 992-999.
6. **Aggarwal R**, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. ***Ann Surg*,** 2007; 246(5): 771-9.
7. **Aggarwal R**, Boza C, Hance J, Leong JJ, Lacy A, Darzi A. Training in laparoscopic Roux-en-Y gastric bypass – an evidence-based approach. ***Obes Surg***, Obes Surg 2007; 17(1): 19-27.

Presentations to Learned Societies

1. Aggarwal R, Moorthy K, Hance J, Grantcharov T, Darzi A. The establishment of a structured virtual reality training curriculum for laparoscopic skills training.
   European Association of Endoscopic Surgeons, Barcelona, 2004.
   Society in Europe for Simulation Applied to Medicine, Stockholm, 2004.
   Society of Laparoendoscopic Surgeons, New York, 2004 (Prize for Best Multi-speciality Paper).
   In *Surg Endosc* 2005; 19 (S1): 49.

2. Grantcharov TP, Aggarwal R, Eriksen JR, Blirup D, Kristiansen V, Darzi A, Funch-Jensen P. A comprehensive virtual reality training program for laparoscopic surgery.
   European Association of Endoscopic Surgeons, Venice, 2005.
   In *Surg Endosc* 2006; 20 (S1): S15.

3. Aggarwal R, Moorthy K, Grantcharov T, Papasavas P, Milland T, Dosis A, Bello F, Darzi A. Dexterity analysis for the assessment of laparoscopic procedures in the operating room.
   Association of Surgeons of Great Britain & Ireland, Glasgow, 2005.
   Society of American Gastrointestinal & Endoscopic Surgeons, Miami, 2005.
   In *Br J Surg* 2005; 91 (S1): 52.
   In *Surg Endosc* 2005; 19 (S1): 243.

4. Aggarwal R, Grantcharov T, Moorthy K, Papasavas P, Milland T, Sarker S, Darzi A. A Reliability analysis of video-based rating scales for technical skills assessments in laparoscopic surgery.
   Society of American Gastrointestinal & Endoscopic Surgeons, Miami, 2005.
   In *Surg Endosc* 2005; 19 (S1): 286.

5. Aggarwal R, Boza C, Hance J, Lacy A' Darzi A. Objective assessment of technical skills for advanced laparoscopic procedures.
   European Association of Endoscopic Surgeons, Venice, 2005.
   In *Surg Endosc* 2006; 20 (S1): S102.

6. Aggarwal R, Ward J, Balasundaram I, Sains P, Darzi A. A proficiency-based laparoscopic virtual reality training curriculum leads to a shortening of the learning curve on real procedures.
   Association of Surgeons of Great Britain & Ireland, Edinburgh, 2006.
   European Association of Endoscopic Surgeons, Berlin, 2006.
   Society of Gastrointestinal & Endoscopic Surgeons, Las Vegas, 2007.
   In *Surg Endosc* 2007; 21(S1): S90.
   In *Surg Endosc* 2007; 21(S1): S311.

7. Boza C, Aggarwal R, Darzi A, Lacy A. A high fidelity porcine model for training in laparoscopic Roux-en-Y gastric bypass.
   European Association of Endoscopic Surgeons, Venice, 2005.
   In *Surg Endosc* 2006; 20 (S1): S217.

8. Aggarwal R, Boza C, Hance J, Lacy A , Darzi A. Skills acquisition for laparoscopic gastric bypass in the training laboratory – an innovative approach.
   International Federation for the Surgery of Obesity, Maastricht, 2005 (2nd Best Poster).
   In *Obes Surg* 2005; 15 (7): 969.

Abstract of Thesis

Medical education is undergoing a paradigm shift, from the traditional experience-based model to a program that requires documentation of proficiency. Technological advances in health care, the development of day-case surgery, and the setting of quality-assurance targets have led to a striking reduction in training opportunities for young doctors. It is no longer acceptable, or appropriate, for students at any level of training to practice new skills on patients, even if they have a patient's explicit consent.

At present, simulation-based training is a prerequisite for all high-reliability organizations (e.g., in the airline, nuclear, and oil industries) yet remains a niche player in medical education. The development of validated training curricula for technical skills necessitates structured training programs which are underpinned by objective methods of assessment. This is especially so for new technologies where the potential for medical error at the early part of the learning curve is greater.

Within this thesis, a methodology for technical skills training and assessment has been produced and applied to basic and advanced laparoscopic training curricula, utilising synthetic and virtual training devices, together with assessment based upon motion analysis, video-based rating scales and in-built measures from virtual reality simulators. Assessment parameters have been validated for use in their particular domain through demonstration of superior performance of experienced versus novice operators. Furthermore, skills of novice subjects have

been shown to improve through repeated training schedules, toward that of experienced operators.

The thesis has sought to elucidate the effectiveness of virtual reality simulation-based training to shorten the learning curve on real procedures, in a similar manner to the transfer-effectiveness ratio of flight simulation. Finally, the development of an evidence-based curriculum for technical skills acquisition of an advanced laparoscopic surgical procedure has been defined, and tested. The application of such curricula is the next stage of work to be completed.

With the development of proficiency-based training and practice comes the need to set benchmarks of achievement in skills and behaviours during pre-specified tasks. Such criteria can be used not only to confirm the completion of a particular training module with the attainment of an appropriate level of technical proficiency, but also to reconfirm the soundness of the skills that have been acquired. It is believed that in the future, expertise rather than experience will underlie competency-based practice and specialty certification.

Aims of the Thesis

The individual aims of this thesis were to:

1. Develop a standardized and universal methodology for development of curricula for the acquisition of technical surgical skills, based upon motor learning theories, task analysis and transfer of skill to the operative environment.

2. Define a training curriculum on a virtual reality simulator for laparoscopic skills acquisition, which is graded in a stepwise manner, and utilizes valid parameters for assessment of proficiency.

3. Investigate the feasibility, validity and reliability of laparoscopic skills assessment in the operating theatre which utilizes motion tracking and video-based analysis to assess technical performance.

4. Prove the effectiveness of a virtual reality training curriculum in improving laparoscopic skill on real procedures, in terms of transfer of skill to real procedures, evaluation of the learning curves between simulation-trained and control groups, and measurement of the transfer-effectiveness ratio of the simulator-based curriculum to real procedures.

5. Construct an evidence-based training program for an advanced laparoscopic surgical procedure, which is task and proficiency-based,

though based upon cadaveric porcine models for assessment of technical skill.

Chapter 1

LITERATURE REVIEW

1. 1 Introduction

The first reported laparoscopic cholecystectomy was carried by Philippe Mouret in France in 1987 (Mouret, 1989), and within five years it was established as a feasible alternative to open cholecystectomy (Dubois et al., 1989;Reddick and Olsen, 1989;Dubois et al., 1990;Perissat et al., 1990;Cuschieri et al., 1991;Olsen, 1991;Voyles et al., 1991). Patients accrued the benefits of smaller incisions, a shorter hospital stay, decreased post-operative pain and could resume normal activities within a week (Cuschieri et al., 1991;Gadacz and Talamini, 1991;Peters et al., 1991;Vitale et al., 1991;Soper et al., 1992;Vander Velpen et al., 1993). It was suggested that this approach could lead to the elimination of complications of gallstones and cure acute cholecystitis (Reddick and Olsen, 1989;Dubois et al., 1990;Reddick et al., 1991;Flowers et al., 1991).

Initial reports of success, the enthusiasm of many surgeons and an increasingly competitive healthcare market led to a large number of operators attempting the new technique. However, doubts soon surfaced about its safety, and the qualifications of those performing the procedures (Berci, 1990;Cuschieri et al., 1990;Reddick, 1990;Dent et al., 1991;Perissat and Vitale, 1991;Cuschieri, 1995). This prompted the surgical community to reconsider the training strategy in laparoscopic surgery.

Societies and regulating bodies such as the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and European Association of Endoscopic

Surgeons (EAES) stipulated minimum requirements for those performing laparoscopic surgery, with an emphasis on training both in and outside the operating theatre (Tompkins, 1990;Berci and Sackier, 1991;Dent, 1991;Society of American Gastrointestinal Surgeons, 1991;European Association of Endoscopic Surgeons, 1994). Skills courses and drills were introduced to teach basic psychomotor skills. As the number of surgeons performing the procedure increased, more opportunities were available for novice surgeons to assist. Laparoscopic cholecystectomy gradually became a safe and effective alternative to the open procedure for a diseased gallbladder (Cuschieri et al., 1991;1991;Schirmer et al., 1992;Dunn et al., 1994).

Initial training of psychomotor skills could successfully commence in the skills laboratory free from the pressures of complication or theatre time. Sophisticated models made from cardboard boxes were developed, now known as box trainers, and more recently virtual reality simulators have led to an expansion in the field of possibilities. Not only can the surgeon practice in virtual reality, but also receive objective and instant feedback on their performance.

This chapter introduces the evolution of training in laparoscopic surgery, from box trainers using innate models to the introduction of virtual reality simulations. It reviews the evidence for simulation-based training in laparoscopic surgery, and emphasises the importance of objective assessment of skills during training. The role of different methods to assess laparoscopic skill, including virtual reality simulators, is also discussed.

1.2 Traditional Training in Surgery

Training in surgery traditionally takes the role of the apprenticeship model with a master craftsman teaching skills to a student. The apprentice practices on real cases and gradually learns to perform the procedure without supervision. Having mastered the operation, the trainee becomes an instructor and the cycle repeats.

This model is the basis for the surgical residency system introduced by William Halstead almost a century ago (Halsted, 1904;Barnes et al., 1989). It is an effective time-tested model, having been successful in transferring skills and knowledge from one generation to the next. However, there is reliance upon chance for educational opportunities, being underpinned by the need for a patient with a specific disease to present to achieve a particular educational goal. Some trainees may also view the operating room as a hostile, stressful and uncomfortable learning environment (Wetzel *et al.*, 2006). Teaching a surgeon to perform a procedure in the operating theatre adds time to the operation, which leads to increased costs. Bridges and Diamond estimated the cost of using operating room time for training to be $48,000 per graduating resident (Bridges and Diamond, 1999;Babineau *et al.*, 2004).

The introduction of shortened training programmes, the European Working Time Directive, lengthening waiting lists and trends toward ambulatory surgery have led to a reduction in surgical training time by about two-thirds (Working group on specialist medical training, 1993;Bulstrode and Holsgrove, 1996;Sharp and

Wellwood, 1996;Pickersgill, 2001;Aggarwal et al., 2004a). Patients too have greater expectations and are reluctant to have their procedures performed by trainee surgeons. A suitable adjunct is to provide initial surgical training outside the operating theatre. In the past young surgeons practised their craft, and were examined on, cadavers in the post-mortem room (1943). However, this is costly, of limited availability and considered to be unethical (1947).

In the 1980s, craft workshops were introduced at the Royal College of Surgeons of England (Bevan, 1986;Stotter et al., 1986;Bevan, 1997). They utilised pig viscera to practice complex vascular and gastrointestinal anastomoses outside the confines of the operating theatre. The trainees were also able to open the structure and assess the quality of their skills. These courses remain popular to this day. However, pig viscera are not readily available, require specialist storage with appropriate Health and Safety approval, and have a concomitant risk of transmission of infection (Hamdorf and Hall, 2000).

In response to this, innate synthetic models were developed (Greenhalgh et al., 1987;Hill and Kiff, 1990). One of these was an abdominal wall jig to teach techniques of abdominal wound closure and stoma formation (Hill and Kiff, 1990). A plastic box contained an inflated balloon overlain with a simulated abdominal wall constructed from layered foam. The student could open and close the simulated abdomen and again inspect suture placement at the end.

Not only was the trainee surgeon able to perform surgical tasks without risks of complication, it was also possible to assess competence at these tasks. This enabled the surgeon to attend the operating theatre armed with basic surgical skills and a prior knowledge of complex surgical manoeuvres. This was the first step in surgical training moving away from the traditional 'see one, do one, teach one' approach.

1.3 The Advent of Laparoscopic Cholecystectomy

Prior interest in diagnostic laparoscopy had been poor, but a new era was evident with the introduction of laparoscopic cholecystectomy. This revolutionary surgical technique generated considerable interest, described as 'most dramatic change in surgery since the introduction of anaesthesia' (Royston *et al.*, 1994). Its introduction however, was haphazard and uncontrolled with up to one-third of surgeons being completely self-taught (Cuschieri, 1992). This was due to the fact that there were few trainers and no training programmes available to teach new skills to experienced surgeons. Most surgical specialists had to learn the procedure through trial and error on their patients. This unfortunately led to a sizeable number of complications (Ponsky, 1991;Zucker et al., 1991;Macintyre and Wilson, 1993), Cuschieri describing this expansion in laparoscopic surgery as 'the biggest un-audited free-for-all in the history of surgery' (Cuschieri, 1995).

Experienced biliary surgeons possessed expert skills in open cholecystectomy, and believed the same skills were required for the laparoscopic procedure. However, it has been shown by Figert et al. that there is a lack of transfer of skills from open to laparoscopic surgery (Figert *et al.*, 2001). A comparison of interns with limited open and laparoscopic surgical experience with two groups of residents, one consisting of junior residents with recent and ongoing open and laparoscopic experience, and the other of senior residents with limited laparoscopic though ongoing open surgical experience was performed. The three groups were assessed at intracorporeal knot tying on a box trainer. There were no significant differences between interns and senior residents, though junior residents were significantly better than both groups. The results of this study reflect the need for specific training in laparoscopic surgery for all surgeons new to the procedure, irrespective of their previous open surgical experience.

The skills necessary to perform laparoscopic surgery are different from those of open surgery, being more allied to endoscopy rather than traditional laparotomy (figure 1a). The surgeon enters the peritoneal cavity using a blind technique, use instruments over 12 inches long with only their tips visible, and get accustomed to the fulcrum effect. Procedures are performed by viewing a two-dimensional video image on a screen up to five feet away, with limited tactile feedback (Reinhardt-Rutland and Gallagher, 1996;Gallagher et al., 1998;Crothers et al., 1999).

Figure 1a – Laparoscopic surgery

The problem of learning laparoscopic surgery lay in the traditional methods of acquiring surgical skill. The apprenticeship model had served the surgical community since the beginning of the twentieth century, but this was the first time a completely new technique had been introduced to the entire general surgical population. Very few surgeons had seen a laparoscopic cholecystectomy, let alone able to perform one. Experienced biliary surgeons assumed that their skills at open surgery would transfer to the laparoscopic environment, but they too were novices in this field. As Krummel states, it was necessary for the surgical curriculum to evolve away from the age-old apprenticeship model, and toward the teaching of skills in a systematic and logical fashion by *doing*, rather than through observation (Krummel, 1998). The early phase of the learning curve could be achieved outside the operating room, and once proficiency had been demonstrated the surgeon could proceed to real operations.

1.4 Laparoscopic Training Outside the Operating Room

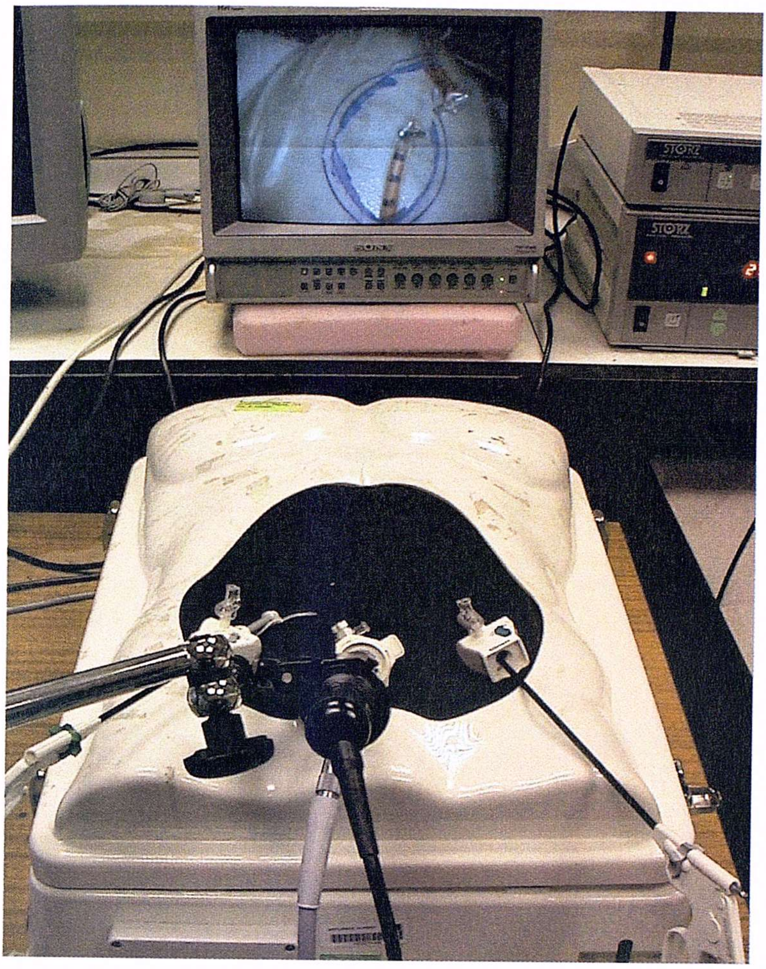
The lack of an expert body of surgeons from whom to learn led to problems in acquiring skills for this new procedure. The group of Berci and Sackier from Los Angeles suggested a modified approach (Berci and Sackier, 1991). Their surgeons scrubbed in with a gynaecologist for five laparoscopic procedures, then performed five laparoscopic procedures on their own patients, for example prior to colon surgery, assisted by the gynaecologist. This led to a safer and more structured introduction to the basics of laparoscopic surgery, but was expensive, time-consuming and difficult to organise.

1.4.1 Skills courses in laparoscopic surgery

It was gradually becoming clear that the apprenticeship model had failed to impart the skills required for laparoscopic surgery. The introduction of laparoscopic courses for general surgeons was an attempt to teach basic laparoscopic skills within a structured and defined curriculum (Darzi, 1996). Courses lasted for two to three days and consisted of didactic lectures with some hands-on simulator experience. The lectures detailed indications and contra-indications to the procedure, with the theory behind establishment of a pneumoperitoneum and its possible complications. Coupled with this were video and live demonstrations of the procedure enabling course participants to recognise and learn important aspects of the operation.

Hands-on simulator experience used models to practice tasks in box trainers, developed from those used in open surgery (Sackier et al., 1991;Majeed et al., 1992;Mughal, 1992) (figure 1b). The models varied from performing simple tasks with different laparoscopic instruments, to teaching the entire procedure using porcine gallbladders (Majeed *et al.*, 1992). At their simplest, candidates were asked to build towers of blocks, or transfer objects between their two instruments, enabling the operator to acquire basic psychomotor skills. Other training models for laparoscopic cholecystectomy deconstructed the procedure into a number of simulated tasks, each of which could be practiced separately.

Figure 1b – A laparoscopic video-box trainer device



Derossis et al. were the first to develop a series of seven laparoscopic tasks (peg transfers, pattern cutting, clip and divide, endolooping, mesh placement and fixation, suturing with intracorporeal or extracorporeal knots) for laparoscopic skills training (Derossis *et al.*, 1998). These were made from simple synthetic objects, and performance was measured using a scoring system rewarding precision and speed. The intention was to develop not only a training module, but also a simulation-based assessment system to identify laparoscopic skill prior to performing procedures within the operating theatre.

Many courses included performing a laparoscopic cholecystectomy on a cadaveric porcine model. However, these models were criticised for being unrealistic. Some courses in mainland Europe and North America enabled the participant to perform a complete laparoscopic cholecystectomy on an anaesthetised pig (Kirwan *et al.*, 1991). This is a contentious point, the use of live animal models being illegal in the United Kingdom, and indeed it is unclear whether high-fidelity models are necessary for acquisition of basic skills (Byrne, 1994;Clayden, 1994;Grober et al., 2004b;Lamata et al., 2006).

The courses provided an opportunity to acquire skills in a controlled and safe environment, free of the pressures of performing procedures on patients in the operating theatre. The aim was to enable surgeons to acquire basic skills for competence in laparoscopic surgery, with the emphasis on the fact that attendance at such courses was not to be interpreted as a licence of proficiency to perform unsupervised laparoscopic cholecystectomies (Morino *et al.*, 1995).

1.4.2 An introduction to virtual reality systems in laparoscopic surgery

The success of training outside the operating theatre in laparoscopic surgery was evident from the increasing number of laparoscopic courses available, with demand regularly outstripping course availability; indeed courses were not only delivered by Academic Surgical Departments, but also by medical device companies keen to increase the use of laparoscopic tools. Nonetheless, box trainers were criticised for being unrealistic and lacking any form of objective assessment (Darzi *et al.*, 1999). Though individuals could practice basic and procedural skills, it was still a far cry from performing the real procedure in a patient.

At the time of the early 1990s, a more sophisticated and quantitative system using computer simulation was in development (Aggarwal *et al.*, 2004b). Computer simulators had been used to train airline pilots for a number of years, providing realistic simulation with an accurate assessment of performance. Indeed, all military and commercial pilots must train and be certified on a flight simulator specific to the aircraft they fly (Satava, 2001a). The possibility of an analogous situation in surgery was beginning to become a reality.

The term virtual reality was first coined by Jaron Lanier in 1989, referring to '*a computer-generated representation of an environment that allows sensory interaction, thus giving the impression of actually being present*' (Aukstakalnis and Blatner, 1992;Pimentel and Teixeira, 1993). Since the first virtual reality laparoscopic surgical simulator developed in 1994, the Virtual Clinic (Cine-Med, Woodbury, Connecticut, USA) (McGovern and McGovern, 1994), considerable advances in computing power and graphics have resulted in higher fidelity simulators.

MIST-VR (Mentice AB, Gothenburg, Sweden) is a basic task trainer, developed in 1996 in collaboration between a psychologist and surgeon (figures 1c and 1d) (Wilson *et al.*, 1997). Earlier simulators had concentrated computing power upon recreation of a realistic environment – a criticism of previous box-trainer devices. However, MIST-VR was developed to teach the psychomotor skills required for laparoscopic surgery, in an abstract environment. It comprises two standard laparoscopic instruments held together on a frame with position sensing gimbals (figure 1c). These link it to a Pentium PC and movements of the instruments are relayed in real time to a computer monitor. Targets appear randomly on the screen, and are 'grasped' or 'manipulated' (figure 1d), though there is no feeling of touch, or force feedback. The simulator software originally consisted of six tasks (core skills 1), and was followed by a further six tasks (core skills 2), with the possibility of performing each task at one of three levels, i.e. easy, medium or hard. The operator works through these tasks with performance measured by time, error rate, and economy of movement for each hand.

In 2001, LapSim (Surgical Science AB, Gothenburg, Sweden) was marketed as a more 'realistic' task-based laparoscopic simulator (figure 1e) (Larsson, 2001). The tasks indeed are more realistic than MIST-VR, involving structures that are deformable and may bleed. Nonetheless, the tasks remain in the mould of training basic and generic laparoscopic skills such as clipping, cutting and instrument navigation. The hardware platform used is the same as the one for the MIST-VR system, thus there is no force feedback to guide the user.

Figure 1c – Minimally Invasive Surgical Trainer – Virtual Reality (MIST-VR)

Figure 1d – MIST-VR tasks

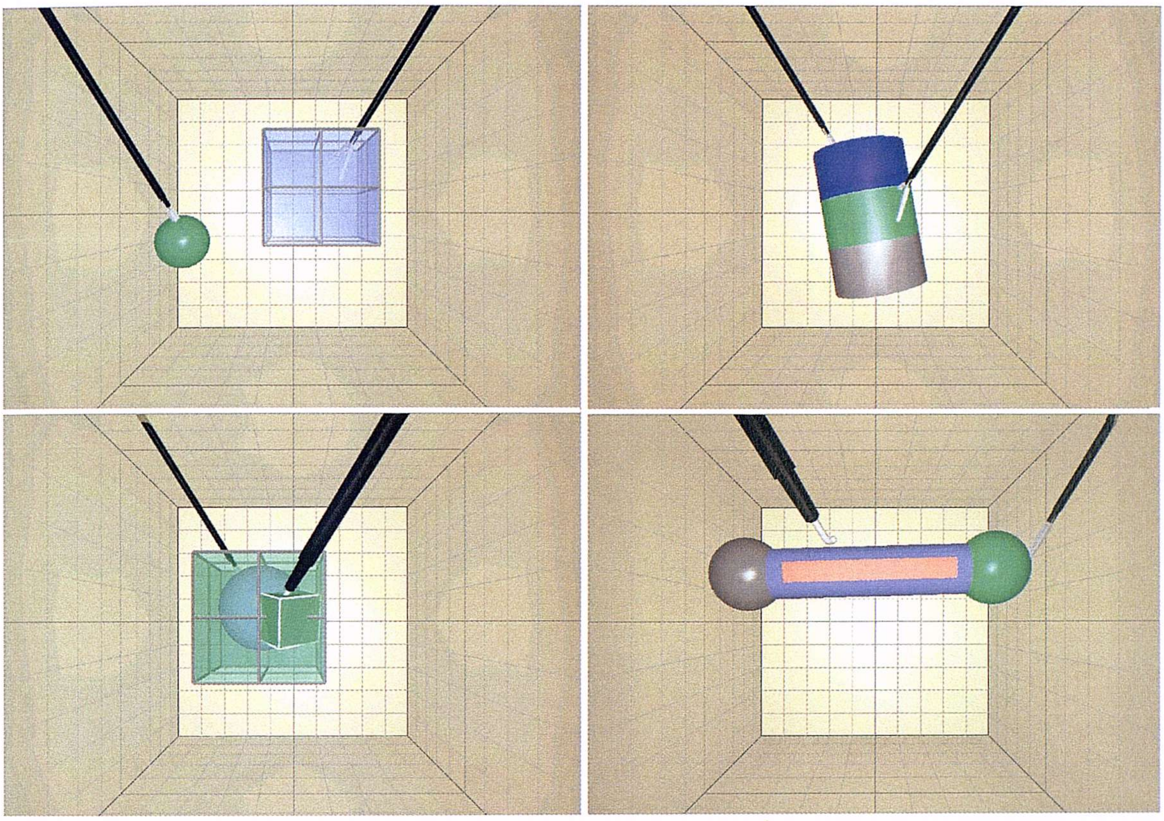
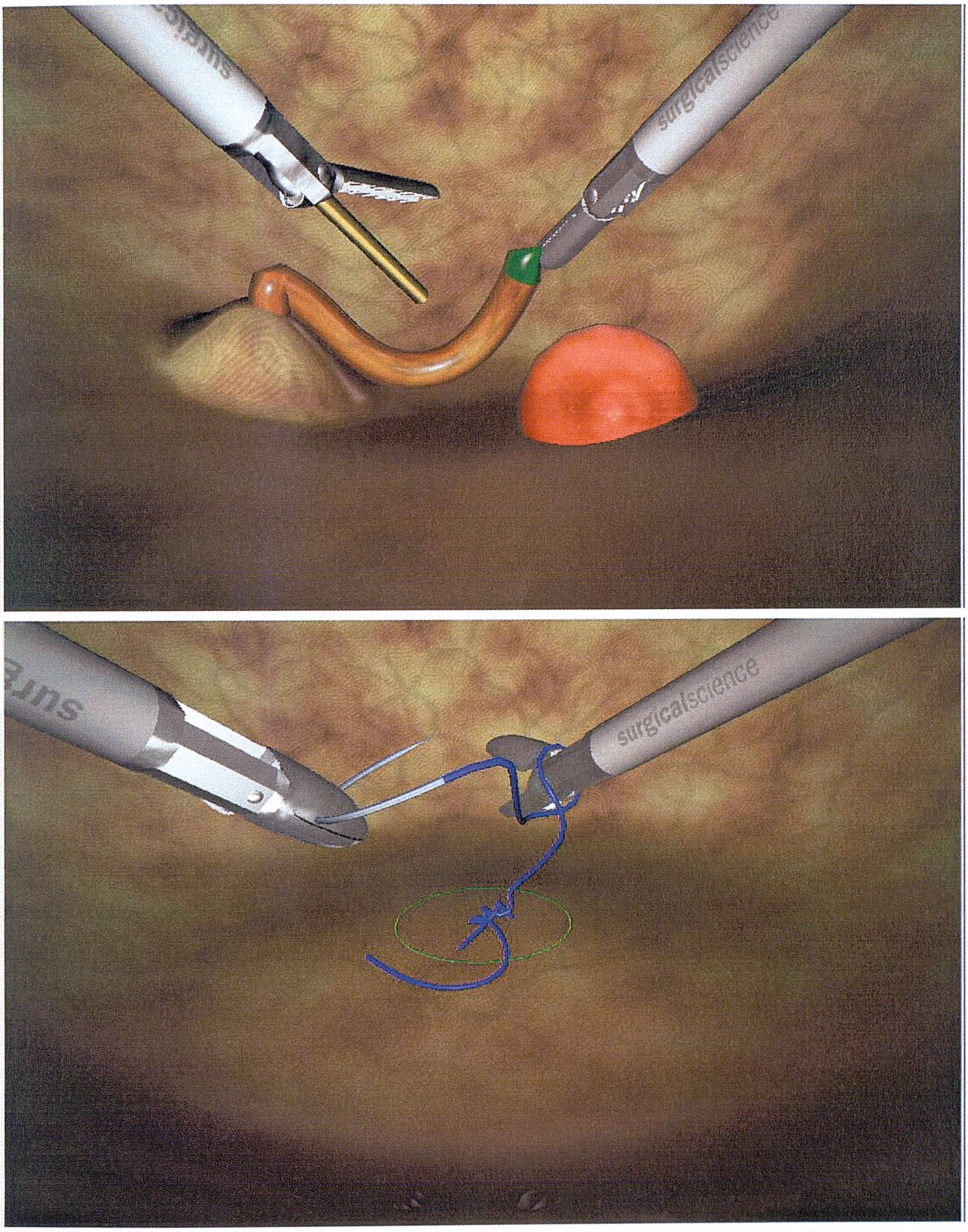


Figure 1e – LapSim tasks



In 2003, the Xitact LS500 (Xitact SA, Morges, Switzerland) laparoscopy simulator, was hailed as one of the 'most promising technical novelties in the area of surgical virtual-reality simulation' (Schijven and Jakimowicz, 2003). It comprises procedure-specific tasks such as dissection, clip application and tissue separation, the intended integration of which can lead to the production of a procedural trainer. A significant departure from both MIST-VR and LapSim platforms is the incorporation of a physical object, the 'virtual abdomen' with active force feedback, i.e. liver tissue 'feels' different to gallbladder or fatty tissue. A questionnaire study by the same authors to 120 expert and trainee surgeons regarding face validity (or realism) of Xitact simulator revealed favourable opinion, though its haptic (or force feedback) abilities were doubted (Schijven and Jakimowicz, 2002).

Within the past three years, a greater number of virtual reality laparoscopic simulator devices have become available. These include the Reachin Laparoscopic Trainer (Reachin AB, Stockholm, Sweden), ProMIS Surgical Simulator (Haptica, Dublin, Ireland), LapMentor (Simbionix, Cleveland, Ohio, United States of America), SurgicalSim Education Platform or SEP (SimSurgery, Olso, Norway) and Simendo (DeltaTech, Delft, The Netherlands). All simulators provide the same basic package of basic laparoscopic skills, with the first three packages additionally exhibiting the ability to simulate entire laparoscopic procedures. The ProMIS simulator is a slight departure from virtual reality simulation in that it offers a hybrid approach of a video-box trainer in which real materials and instruments can be placed, with an overlaid 'augmented reality' image of the task.

Though not the focus of this piece of investigative work, it must also be considered that there a number of virtual reality simulators for application in other domains of medicine, i.e. gastrointestinal endoscopy, bronchoscopy, gynaecology, endovascular procedures, arthroscopy, sinus surgery and tele-robotics.

Current computer-based simulation devices enable the user to practice standardised laparoscopic tasks repeatedly, with instant objective feedback of performance. The simulators are portable, use standard computer equipment and are commercially available. With graded exercises at different skill levels, it is possible that they can be used as the basis for a structured training program. The application of such a training program should then transfer to improved performance in the clinical domain. This is the nature of the research which has been carried out within this piece of work.

1.4.3 Learning curves on laparoscopic trainers

Box and virtual reality trainers can teach laparoscopic skill, but one question that has eluded the surgical community is their efficiency in transferring skills to the real procedure. A pilot learns to fly on the simulator prior to embarking on a real flight, reducing the time required to become competent in the air. Similarly it is possible that laparoscopic simulators can reduce the amount of time required to train in the operating theatre. This may be proven indirectly by performing learning curve studies with experts and novice surgeons. The hypothesis states that if the simulator accurately mimics skills required of expert laparoscopic surgeons, then the learning curves for this group should be flat and novice surgeons should have steep learning curves. Repeated practice should lead to the scores for novice surgeons approaching those of expert surgeons.

A series of studies have employed this hypothesis, some of which are reviewed in this section. Smith et al. devised an innate laparoscopic simulator containing ten posts of varying heights, placed within a video-box trainer (Smith *et al.*, 2001). Ten non-surgeons were asked to touch each post in a sequential order, each participant performing ten repetitions. The laparoscopic instruments were connected to sensors recording their position in space in three dimensions. This enabled a calculation of the total path length per activity, and was compared to ideal path length between the ten posts, i.e. 'as the crow flies'. Hence not only was time, but accuracy of movement also a measure. Subjects improved their time taken by 70% in the first three repetitions, but not thereafter. Accuracy of task completion improved by 40% over the ten repetitions, but these measurements did not appear to have reached a steady state. This implies that time taken is only one factor in skills assessment, the learning curve for operator speed being shorter than that for operator accuracy.

Gallagher and Satava reported data on 12 experienced (>50 procedures), 12 inexperienced (<10 procedures) and 12 novice laparoscopic surgeons who completed ten trials of the first six tasks (core skills 1) on MIST-VR (Gallagher and Satava, 2002). Results for all six tasks were summed together, with trials analysed from one through to ten. Upon analysis of time taken, trial number one revealed significant differences between the three groups, and as expected the experienced group were the fastest and most consistent. The novice group took longest to complete the tasks and showed the greatest variability in their scores. Time taken for all groups had plateaued by trial five, with the improvements greatest for the novice and inexperienced groups. Similar results were reproduced for the other four variables recorded, namely economy of movement for each hand, economy of diathermy and error scores.

The results of this study also emphasise the importance of variability of scores when considering skill acquisition in laparoscopic surgery. Experienced surgeons perform to a higher standard, but as a group are also more consistent in their performance. With repetition on the simulator, a group of novices can reduce the variability in their scores. Training on this simulator produces a group of novices who can all attain a standard similar to that of the experienced surgeon. Grantcharov et al. (Grantcharov *et al.*, 2003) and Chaudhry et al. (Chaudhry *et al.*, 1999) have performed similar studies on the MIST-VR simulator with comparable results.

Sherman et al. published a paper in 2005 on the learning curve for the acquisition of laparoscopic skills on the LapSim virtual reality simulator (Sherman et al., 2005). Three groups of subjects (expert, junior, and naïve) underwent 12 repetitions of three tasks on the simulator: grasping, cutting, and clipping. The grasping task requires the subject to pick up virtual objects appearing on the screen and then place them in a target, releasing the object when the program recognizes proper positioning of the object in a simulated three-dimensional target. The cutting task involves grasping a vessel with the right instrument and applying cautery with the left instrument. A foot pedal, which is attached to the system, is pressed down to simulate burning of the tissue. Once the vessel is transected, the free fragment is placed in the three-dimensional target previously described. The clipping task involves grasping the vessel with one instrument while applying clips to a predetermined area on the vessel. This manoeuvre is repeated at the other end of the vessel, with instruments reversed. Once the clips are adequately placed, the vessel is transected using scissors. In a similar manner to the aforementioned studies, all three groups improved significantly from baseline to final sessions. The novice group had significantly higher time and error scores at end of training as well as at commencement, perhaps suggesting that the learning curve for this simulator to be longer than that for MIST-VR. Nonetheless, another study has questioned whether even 30 repetitions on the MIST-VR simulator (as opposed to the unsubstantiated standard of 10 repetitions) are enough? (Brunner *et al.*, 2004)

These studies provided the initial evidence for the role of virtual reality simulation to improve technical skills (on the simulator) through repeated practice. But did this actually lead to improved laparoscopic skill, or did the subjects just 'learn the computer game on the simulator'?

1.4.4. Transfer of laparoscopic skill from simple to complex tasks

A further criticism of laparoscopic trainers is that they are abstract, with tasks being too simple and not related to real procedures. Rosser et al. described a structured training method to enhance laparoscopic surgical skills with 150 trainee surgeons performing a series of three standardised simple laparoscopic drills, on a box simulator, ten times each (Rosser *et al.*, 1997). At the end of each drill, their skill at performing the advanced laparoscopic task of an intracorporeal stitch was assessed by time taken. After only one series of tasks, the time taken to complete an intracorporeal stitch improved significantly, with further significant improvement at the end of ten repetitions. An analogous experiment using MIST-VR as the training tool revealed significant improvements in knot-tying times after a five-day training period (Kothari *et al.*, 2002).

The evidence for surgical simulation, either box-trainer or virtual reality based is perhaps compelling. Repeated practice leads to performance variables similar to those of experienced surgeons, and the acquisition of basic skills can also be internalised by trainees with concomitant improvements during a complex task. However, the ultimate question is whether this departure from the standard apprenticeship mode of training can equate to improved skill in the actual operating theatre? Initial thoughts were varied, leading to preliminary experimentation within the animal operating theatre.

1.4.5. Transfer from laparoscopic trainer to animal model

Twelve general surgery residents performed basic laparoscopic tasks on an inanimate laparoscopic trainer and an *in vivo* porcine model (Fried *et al.*, 1999). These skills were similar to the Rosser tasks, and included transferring, cutting, clipping, placement of a ligating loop, mesh placement, and suturing with an intracorporeal and extracorporeal knot (Rosser *et al.*, 1997). Half the group then received five, weekly one-hour practice sessions on the inanimate, or box-trainer model. The two groups were retested on both models and it was found that everyone had significantly improved their performance in vivo. Notably, the practice group had a significantly greater improvement than the control group, suggesting that practice on a laparoscopic trainer had improved skill in vivo. However, the in vivo porcine model did not assess a real procedure, rather the same abstract skills as the inanimate trainer.

A similarly designed study assessed the training potential of the LapSim virtual reality simulator, again with abstract in vivo tasks on the porcine model (Hyltander *et al.*, 2002). 24 medical students were recruited, 12 of them undergoing five, weekly two-hour practice sessions on the LapSim simulator. The in vivo assessment was performed at the end of the training period, revealing significantly better results in the trained than control group on the basis of time, and evaluation scores from four blinded expert surgeons.

In answer to the criticisms of the aforementioned studies, Grantcharov et al. assessed the skills of 14 inexperienced laparoscopic surgeons (<10 procedures) at performing a complete laparoscopic cholecystectomy on an anaesthetised pig. Their scores were compared with their performance on the MIST-VR simulator (Grantcharov *et al.*, 2001). Though assessment of the in vivo procedure was performed using an un-validated operative rating scale, there were significant correlations ($p<0.05$) between scores obtained on the in vivo procedure and some of the virtual tasks. However, this was not strictly a study to prove the effect of training on the simulator; rather it demonstrated a correlation between performance on the simulator and on a porcine model. The confirmation for the use of laparoscopic trainers in a surgical training curriculum must lead to an improvement in performing a laparoscopic operation on human patients. This was the next step to ensure the growth of virtual reality simulation for laparoscopic skills training.

1.4.6 Transfer of skill from laparoscopic trainer to human patient

Scott et al. recruited 22 second and third-year residents to perform a laparoscopic cholecystectomy on a live human patient in the operating room (Scott *et al.*, 2000a). A validated global rating scale was used to assess operative skill, the operation carried out with a one- or two-handed laparoscopic technique, dependent upon the senior surgeon's preference. Subjects were randomised into two groups, and there was no difference in baseline scores between the two groups. Nine residents randomised to simulator-training underwent 30 minute session per day for ten days on five laparoscopic box trainer tasks; the remainder underwent no formal training.

The training curriculum was based upon five established laparoscopic drills suitable for novice surgeons that could be performed on a box-trainer, i.e. checkerboard, bean drop, running string, block move, and suture foam. The checkerboard drill involves arranging 16 metal letters and numbers in the appropriate squares on a flat surface. The bean drop drill consists of individually grasping five beans and moving the beans 15cm to place them in a 1-cm hole at the top of an elevated cup. The dominant hand is used to grasp the beans while the non-dominant hand moves the laparoscope to provide adequate visualization during the procedure. The running string drill mimics running bowel; two graspers are used to run a 140-cm string from one end to the other, grasping the string only at coloured sections marked at 12-cm intervals. The block move drill consists of individually lifting four blocks using a curved needle (held in a grasper) to hook a metal loop on the top of each block. The dominant hand manipulates the grasper to move the blocks 15cm and to lower them onto a designated space on a flat surface. The non-dominant hand moves the laparoscope to provide adequate visualization during the procedure. The Suture Foam drill consists of using an Endostitch device (United States Surgical Corporation, Norwalk, CT) to suture two foam organs together and tie a single intracorporeal square knot.

In the post-assessment, the trained group improved to a significantly greater extent than the control group on four of the eight assessment criteria (p=0.005 to 0.035), with non-significant improvements on the remaining criteria (p=0.058 to 0.100). This was the first study to definitively relate improvement in skill level on inanimate models with improved skill level in live human operations. Scott et al. had shown that skills acquired on a laparoscopic simulator are transferable to the operating room.

In 2002 Seymour et al. published the first study to assess the training potential of a virtual reality simulator in this manner. It involved sixteen surgical residents of varying experience (Post Graduate Year 1-4), randomised to training on the simulator, or control (Seymour *et al.*, 2002). Baseline abilities were assessed using commercially available psychomotor tests, with no difference between the two groups. The former group trained solely on the diathermy task of MIST-VR until they attained a previously specified expert score. For this task the trainee is required to hold a virtual sphere within the centre of a slightly larger virtual cube, and 'diathermy' smaller cubes from the side of the sphere. The aim is to test two-handed co-ordination, accuracy and task efficiency.

Then, all subjects performed a human laparoscopic cholecystectomy, scored by two independent observers using a checklist for gallbladder dissection. The trained group were found to dissect the gallbladder 29% faster, were nine times more likely to make progress and five times less likely to make errors. However, there was no evaluation of baseline laparoscopic skill prior to the study, merely the non-specific and un-validated tests of psychomotor ability.

There was need for a more conclusive study, with similar design to that of Scott et al (Scott *et al.*, 2000a). In 2004, Grantcharov et al. published a study in which they recruited 16 laparoscopic novices (<10 procedures) and divided them into two groups of eight, a MIST-VR trained group and a control group (Grantcharov *et al.*, 2004). The MIST-VR trained group had ten sessions on the simulator, each session encompassing the completion of six basic tasks, i.e. core skills 1. The six tasks are of progressive complexity and are designed to simulate the techniques used during laparoscopic cholecystectomy. All tasks begin with bilateral movements to touch a virtual sphere with the tips of the virtual instruments. For task 1 the trainee is required to grasp a virtual sphere and place it in a virtual box. In task 2 the virtual sphere is grasped, transferred between instruments and then placed in the box. Task 3 consists of grasping alternately the segments of a virtual pipe. Task 4 requires the trainee to grasp the virtual sphere, touch it with the tip of the other instrument, withdraw and reinsert this instrument, and once more touch the sphere. In task 5, once the virtual sphere has been grasped, three plates appear on the surface of the sphere, 90° apart; these are then touched by the other instrument and, using the pedal, virtually diathermied away. Task 6 combines the actions of tasks 4 and 5 with the aim of diathermying the plates while holding the sphere in the virtual box.

All subjects underwent a pre-training and post-training assessment of laparoscopic skill by performing a human cholecystectomy which was videotaped. Assessment of the operations was performed using a modified global rating scale. The trained group performed the post-assessment significantly faster than the control group ($p=0.021$), and had greater improvement in their error ($p=0.003$) and economy of movement ($p=0.003$) scores.

Though performed with small numbers of subjects, these studies have shown transfer of laparoscopic skills training from box-trainer and virtual reality systems to real procedures. Though likely to be a more cost-effective method of imparting surgical skill, it is necessary to determine the role of box-trainer models and virtual reality simulations within the training curriculum – is one superior to the other?

1.4.7 Comparison of standard and virtual laparoscopic trainers

Jordan et al. sought to tackle this question with a preliminary study comparing virtual reality and box-trainer practice with a control group (Jordan *et al.*, 2001). 32 laparoscopic novices were randomised to one of four groups. One group trained on MIST-VR, one group on a 'U-type' laparoscopic maze-tracking task on an innate trainer, a third group on a 'Z-type' laparoscopic maze-tracking task, and a control group who received no training. Though the pre- and post-training assessment task favoured the box-trainer groups inasmuch as being a laparoscopic cutting task performed on a box trainer, subjects who trained on MIST-VR made a significantly greater number of correct incisions (P<0.0001) and fewer incorrect incisions (p<0.0001) than all other groups.

A further study performed by Torkington et al. compared training on the box simulator with MIST-VR (Torkington *et al.*, 2001). Thirty right-handed clinical medical students with no previous exposure to laparoscopic surgery were assessed with the use of a box-trainer which involved the subjects to grasp five sutures sequentially with a grasping forceps in their left hand and to cut each one with a laparoscopic scissors held in the right hand. Randomisation was into three equivalently sized groups: group one received no further intervention; group two received training of both hands using the MIST-VR, completing 10 repetitions of core skills 1 training tasks; group three received one hour of training using standardized minimal-access training drills, as utilized in the Royal College of Surgeons Basic Surgical Skills Course. These include placing chick peas on golf tees, passing matchsticks through hoops, cutting out shapes drawn on surgical gloves, stacking sugar cubes, and unwrapping confectionery using laparoscopic grasping forceps. The trained groups performed to a significantly higher standard on the post-test, compared to control subjects, though there were no statistical differences between virtual reality and conventional trained groups.

A comparison of training on the LapSim simulator with a box trainer in the acquisition of laparoscopic skill also reported equivalence of skills acquisition (Munz *et al.*, 2004). In this study, 24 novice surgeons were tested to determine their baseline laparoscopic skills and then randomized into the following three groups: LapSim, box trainer, and no training (control). After 3 weekly training sessions lasting 30 minutes each, all subjects were reassessed on a previously validated box-trainer task, with box-trainer and virtual-training groups achieving significantly better scores than the control group. The study authors conclude that 'at present, it seems that the effectiveness of these computer-based systems is questionable, which is probably why proposals to integrate VR into the surgical training curriculum have been met with some scepticism.'

It is also possible that the reason for lack of a superior effect of virtual-reality versus traditional box training is because of the assessment tasks were all performed on a box-trainer. Though the tasks were different, there is a definite bias toward this mode of training. This criticism has been addressed in an elegantly designed study, published by Hamilton et al. in 2002. They randomised 31 first, and 19 second, year residents to 10 half-hour sessions on a box trainer or MIST-VR (Hamilton *et al.*, 2002). The subjects underwent baseline and post-training skills assessments on *both* the box trainer and MIST-VR simulator, with all subjects achieving significant improvements regardless of which simulator they trained on (p<0.001). Importantly though, the virtual reality training group improved more on the box-trainer post-test tasks (36%) than the box-trainer group improved on the virtual reality post-test tasks (17%). This result was statistically significant (p<0.05). Though not definitive, this study certainly approves the role of both virtual reality and box trainer systems in laparoscopic skills acquisition.

Youngblood et al. recently published a study which compared the effectiveness of box trainer and virtual reality simulation over a control group by assessing the transfer of skills learned on simulators to closely matched surgical tasks in the animal laboratory, i.e. on a different modality to that used for training (Youngblood *et al.*, 2005). Forty six medical student volunteers were randomly assigned to one of three groups: box trainer group (n = 16), LapSim group (n = 17), and control (n = 13). Outcome measures included both time and accuracy scores on three laparoscopic tasks (grasp and place, running the bowel, clip and cut) performed on live anaesthetised pigs, and a global rating of overall performance as judged by four experienced surgeons. The LapSim group performed significantly better than the box-trainer group on 3 of 7 outcomes measures, though the box trainer group did not perform significantly better than the LapSim group on any measure. Of concern though was the result that the box trainer group performed significantly better than the control group only on 1 of 7 outcome measures, and the LapSim group performed significantly better than the Control group on only 2 of 7 measures. No explanation is offered for this conflicting result, though perhaps the period of training was too short.

Despite the acquisition of technical skill, virtual reality simulators possess further advantages such as their portability, versatility and the ability to devise structured, stepwise curricula. All trials are recorded, enabling instant, objective feedback for the trainee, trainer and training establishments. Preparation for a task simply involves a mouse click, rather than having to ensure all relevant instruments and materials are available. Expense is a commonly stated problem, but with continued developments in computer technology, virtual reality simulators should become both cheaper and more technologically advanced.

1.4.8 Toward confirmation of the role of virtual reality simulators in training

Concerning transfer of laparoscopic skills to the operating room, the 19 second year residents in Hamilton's study were pre- and post-assessed at performing a laparoscopic cholecystectomy (Hamilton *et al.*, 2002), using a previously validated global rating scale (Martin *et al.*, 1997). The MIST-VR trained group improved their scores significantly ($p<0.01$), whilst the box trained group achieved non-significant improvements in operative skill. This is in tandem with results of previously mentioned studies, but the fact remains that present evidence for the role of VR simulators in transfer of skills to the operating theatre is still weak. There is a lack of standardisation of all procedures, variation in the operative techniques used, and assessment performed by a variety of observational methods.

A recent systematic review of the literature by Sutherland et al. sought to analyse randomized controlled trials assessing any training technique using at least some elements of surgical simulation (i.e. box trainer or virtual reality), which reported measures of surgical task performance (Sutherland *et al.*, 2006). Thirty randomised controlled trials (RCTs) with 760 participants were able to be included, although the quality of the RCTs was stated to be poor. Computer simulation generally showed better results than no training at all (and than physical trainer/model training in one RCT), but was not convincingly superior to standard training modes. The authors conclude that 'while there may be compelling reasons to reduce reliance on patients, cadavers, and animals for surgical training, none of the methods of simulated training has yet been shown to be better than other forms of surgical training.'

The few currently published studies on the transfer of skill to the operating room are also limited by difficulties in objectively assessing a surgeon's performance in the operating room. Time alone has been shown to be a poor indicator (Darzi *et al.*, 1999). The development of exclusively objective measures of operative skill are important to confirm the role of simulators in laparoscopic surgery, but if successful may also lead to further advances in credentialing and revalidation of all surgeons.

1.5 Assessment of Technical Skills in Laparoscopic Surgery

The introduction of laparoscopic cholecystectomy led to, in 1991, the Society of American Gastrointestinal Endoscopic Surgeons (SAGES) publishing minimum requirements for 'Granting of Privileges for Laparoscopic General Surgery' (Society of American Gastrointestinal Surgeons, 1991). 'Surgeons should have completed their general surgical training with credentials in diagnostic laparoscopy. Training should be by a surgeon experienced in laparoscopic surgery, or on completion of an approved course with hands-on laboratory experience. Furthermore, the applicant's laparoscopic training director should confirm training, experience and observed level of competency'. Similar guidelines were produced by the Society for Surgery to the Alimentary Tract (SSAT) (Dent, 1991), and these models were used by individual hospitals to devise their own privileging criteria.

This was the first time that guidelines (as opposed to the apprenticeship model of training) had been laid down for surgeons prior to performing an operative procedure. Definitions of competency differed amongst institutions, but they all shared common ground (Tompkins, 1990;Dent, 1991;Society of American Gastrointestinal Surgeons, 1991). Competency was based upon the number of procedures performed or time taken, both crude and indirect measures of technical skill, or upon evaluation of the trainee by senior surgeons, a process that is known to be subjective and biased (Elliot and Hickam, 1987;Van Rij et al., 1995;Hanna et al., 1998b;Warf et al., 1999). Professional organisations had recognised the need to objectively assess surgical performance (Jackson B, 1999), though not the methods to accomplish this through analysis of a surgeon's technical skill.

Technical proficiency is a *sine qua non* of safe surgical practice. Assessment of technical skills in laparoscopic surgery consists of dexterity analysis and video-based assessment, and may be used for quality assurance of a surgeon's skill and credentialing for new procedures (Aggarwal *et al.*, 2004b). This may also enable structured progression during training, together with identification of poorly performing trainees who may require remedial action. Objective assessment of skill and feedback of performance to the trainee is crucial to the learning process, with the premise that there can be no learning without objective feedback (Reznick, 1993).

There is a shift in the philosophy of medical education to deliver a competency-based system of achievement and progression. It is not time spent on the ward, in the operating room, or with a tutor which is important, but the attainment of pre-determined goals. In August 2005, the Foundation Programme was introduced in the United Kingdom for all newly qualified doctors (Neville, 2003). The aim is to develop specific and focused learning objectives, with in-built demonstration of clinical competence prior to progression onto specialist or general practice training. Each trainee completes a number of assessments, leading to the development of a *portfolio* of clinical performance. This must be underpinned by objective and reliable modes of assessment.

1.5.1 Fundamental principles of assessment in surgical education

In order for any method of skill assessment to be used with confidence, it must be reliable, valid and feasible (Gallagher *et al.*, 2003a). **Reliability** is a measure of precision of a test, and supposes that if a test were repeated on two separate occasions, with no learning between the two tests, then the results would be identical. It is measured as a ratio from 0 to 1.0, a test with reliability of 0 to 0.5 being of little use, 0.5 to 0.8 being moderately reliable, and those with a ratio of over 0.8 being the most useful.

Validity refers to the concept of whether a test measures what it purports to measure, and may be divided into five parts. *Face validity* refers to whether the model resembles the task it is based upon, and *content validity* questions to what extent it measures surgical skill, and not simply anatomical knowledge. *Construct validity* is a test of whether the model measures what it purports to measure, and is extrapolated from its ability to differentiate between different levels of experience. *Concurrent validity* compares the test to the current gold standard, and *predictive validity* as to whether the test corresponds with future performance in the operating theatre environment.

Over the past decade, considerable developments have been made in the objective assessment of technical skill in surgery. These can be broadly classified as dexterity analysis systems and video based assessments.

1.5.2 Dexterity analysis for laparoscopic skills assessment

Laparoscopic surgery lends itself particularly well to motion analysis as hand movements are confined to the limited movements of the instruments. Three devices for motion tracking of hand movements have been described in the literature.

1.5.2.1 The Skills Assessment Device (SAD)

As previously stated, Smith et al. connected laparoscopic forceps to sensors to map their position in space, which then relayed movements of the instruments to a personal computer (Smith *et al.*, 2001). This enabled calculation of the instrument's total path length, and was compared to the minimal path length to complete the task. This Skills Assessment Device (SAD) measured instrument position at a rate of 40 samples per second, accurate to 0.01mm, results showing the learning curve for operator speed to be shorter than the learning curve for operator accuracy. The authors stated that time 'is a rather crude and gross vehicle for measuring performance' and that training to speed alone may slow the acquisition of precision and accuracy. This sentiment has been echoed by Darzi et al., the developers of a motion tracking device for use in both open and laparoscopic surgical procedures (Darzi *et al.*, 1999).

1.5.2.2 The Imperial College Surgical Assessment Device (ICSAD)

The Imperial College Surgical Assessment Device (ICSAD) has thimble-sized (1cm square) sensors placed on the back of the surgeon's hands, and a 4cm square electromagnetic emitter placed within 100cm of the surgical field (figures 1f and 1g) (Torkington et al., 2001;Datta et al., 2001;Smith et al., 2002;Moorthy K et al., 2003). A commercially available device emits electromagnetic waves (Isotrack II, Polhemus, CT, USA) to track the position of the sensors in x, y and z axes 20 times per second (figures 1f and 1g). Data is analysed in terms of time taken, distance travelled (economy of movement) and total number of movements for each hand. It is also possible to calculate speed of movement for each hand. Previous studies have confirmed the construct validity of ICSAD as a surgical assessment device for laparoscopic procedures, both for simple tasks (Taffinder *et al.*, 1999), and entire procedures such as a laparoscopic cholecystectomy (Smith *et al.*, 2002).

The former study involved 92 subjects of varying laparoscopic experience performing standard tasks in a box trainer (Taffinder *et al.*, 1999). Surgeons performing laparoscopic surgery regularly were significantly more efficient and made fewer movements than occasional laparoscopists ($p<0.02$), who in turn were better than those who had no previous laparoscopic experience at all ($p<0.01$). ICSAD data was reliable, and correlated with scores generated by pre-trained observers watching videotapes of the tasks.

In the latter study, 15 surgeons of varying laparoscopic experience performed three laboratory-based laparoscopic cholecystectomies each on a cadaveric porcine model whilst wearing ICSAD sensors (Smith *et al.*, 2002).

Figure 1f – The Imperial College Surgical Assessment Device (ICSAD)

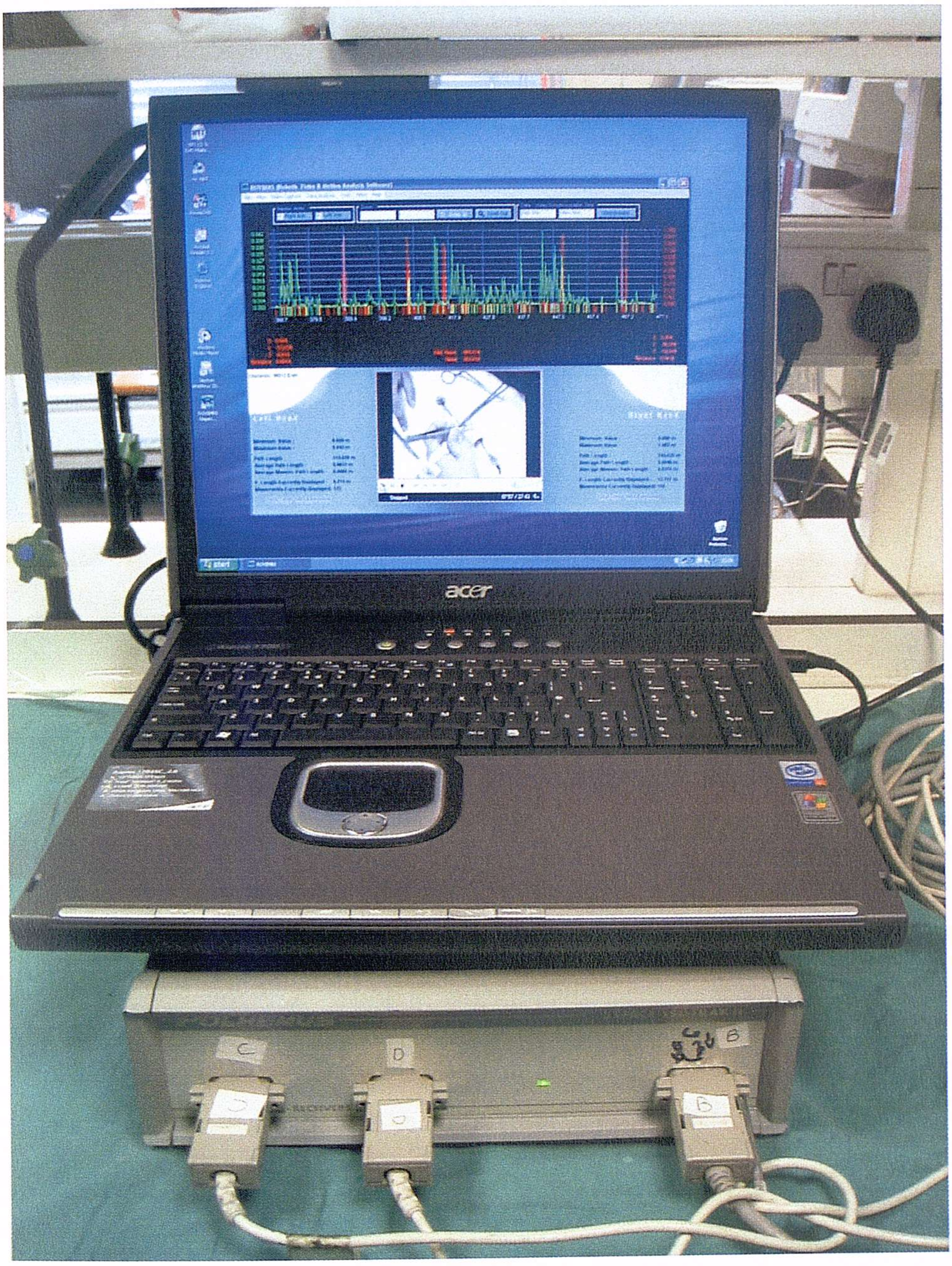
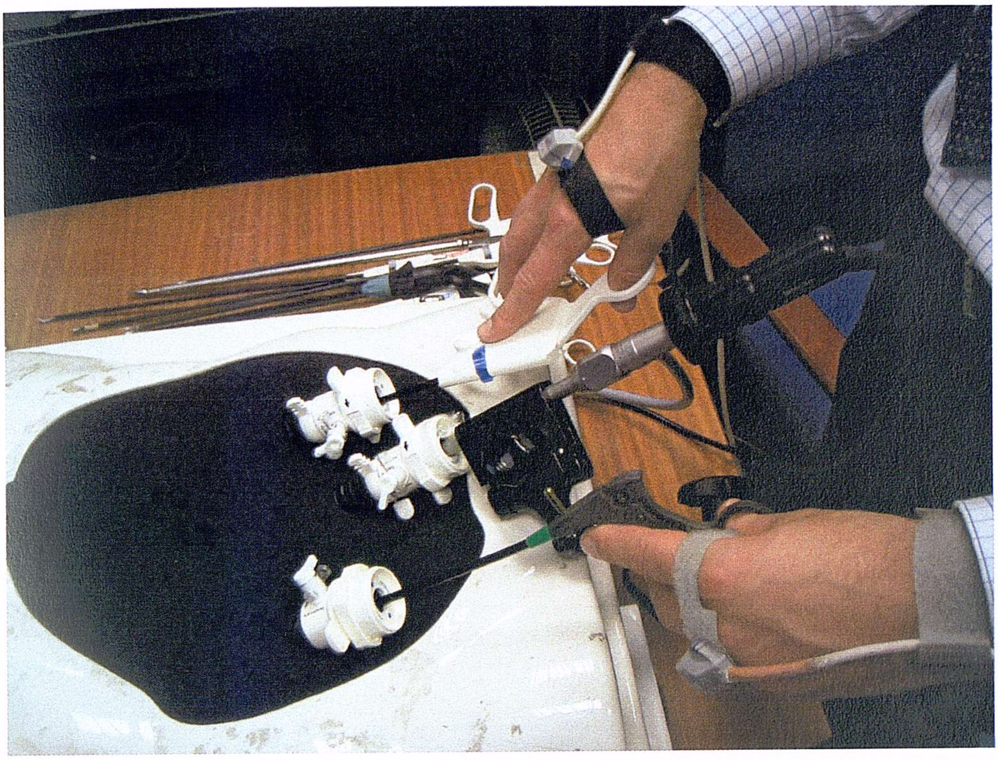


Figure 1g – The Imperial College Surgical Assessment Device (ICSAD) in use

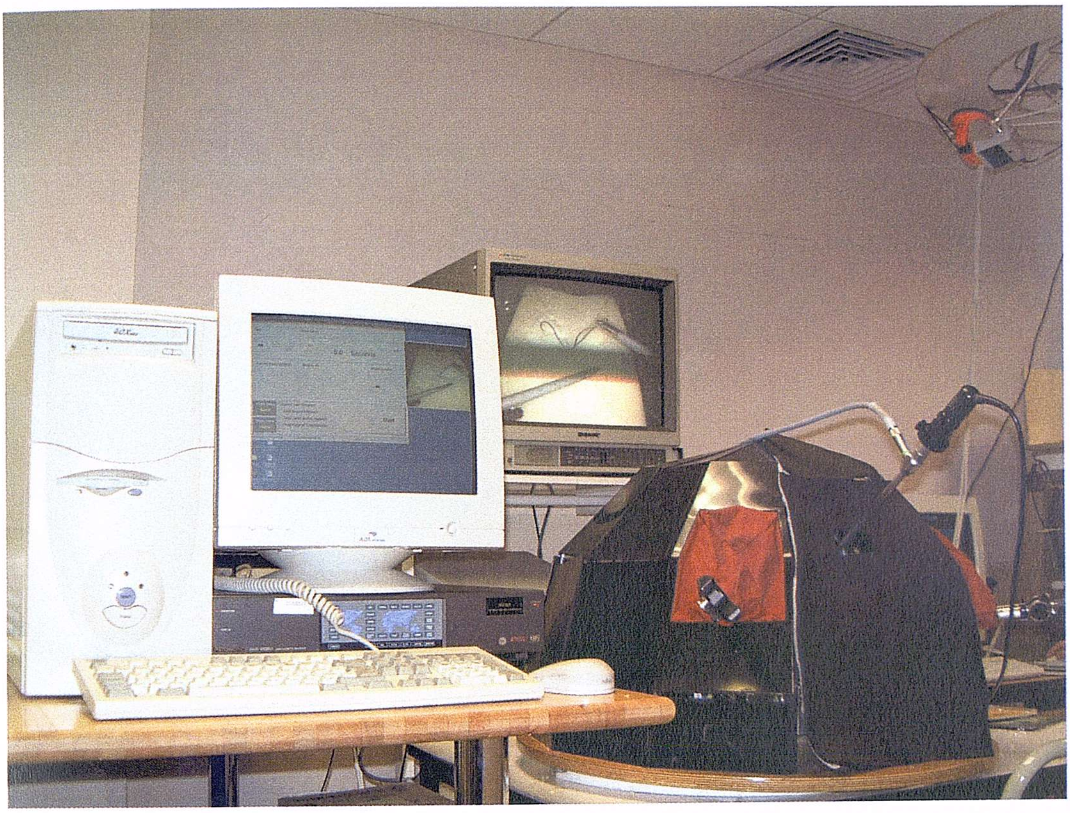


The procedure was standardised, and subjects divided into experienced (>100 laparoscopic cholecystectomies), intermediate (10-100 procedures) and novice (<10 procedures) surgeons. One-way analysis of variance with level of experience as a factor and motion analysis scores as dependents revealed significant differences in time, distance and number of movements for dissection of Calot's triangle and clipping of the duct or artery. Time taken, number and speed of movements were also significant factors in dissection of the gallbladder from the liver bed, whilst parameters for cutting the duct or artery were not significantly different amongst the three groups. These two studies demonstrate experienced laparoscopic surgeons perform the task with a lower path length when compared to inexperienced surgeons, the assumption being that this reflects their significantly higher level of accuracy and greater economy of motion. This tool has been used successfully in a number of research studies for both open and laparoscopic surgical dexterity assessment, within the skills laboratory and in the operating theatre.

1.5.2.3 The Advanced Dundee Endoscopic Psychomotor Tester (ADEPT)

The Cuschieri group, being one of the centres of innovation in the early days of laparoscopic cholecystectomy, developed the Advanced Dundee Endoscopic Psychomotor Tester (ADEPT) (figure 1h). The aim was to develop a research tool to aid definition of the optimal ergonomic conditions during laparoscopic surgery. ADEPT is a computer-controlled device consisting of a dome enclosing a defined workspace, with two standard laparoscopic graspers mounted on a gimble mechanism to measure angular deviations in the x, y and z axes (Hanna *et* *al.*, 1998a). Within the dome is a target plate containing five innate tasks, overlain by a spring-mounted Perspex sheet with apertures of varying shapes and sizes. A standard laparoscope relays the image to a video monitor. Each task involves manipulation of the top plate with one instrument to enable the other instrument to negotiate the task on the back plate through the access hole. Tasks are randomly picked by the computer, for example, number of times a switch must be flicked, or how far a toggle has to be slid along a scale. The system registers time taken, successful task completion, angular path length and instrument error score – a measure of instrument contact with the sides of the front plate holes.

Figure 1h – Advanced Dundee Endoscopic Psychomotor Trainer (ADEPT) – reproduced with the kind permission of Mr. G. Hanna, Imperial College London, England.



20 master surgeons and 20 junior surgeons were tested on the ADEPT system, the former exhibiting significantly lower instrument error rates than the trainees (p=0.007), though no significant differences in execution time or task completion scores were noted (Francis et al., 2002). This study was said to confirm construct validity of ADEPT 'because master surgeons completed the tasks more accurately without sacrificing execution time'. A further study recruited 10 specialist registrars to undergo 10 runs each on ADEPT, each run consisting of 20 tasks (Macmillan and Cuschieri, 1999). Each subject also underwent a blinded clinical assessment by up to four expert surgeons who had operated with, and trained, the subject within the past three years. This was performed on a competency scale from 1 to 4. Overall performance on ADEPT correlated well with assessment of clinical competence (r=0.789), a measure of concurrent validity. There was however no correlation between competence and time taken to perform the tasks (r=0.006). Test-retest reliability of the system produced positive correlations for all three parameters (r=0.60 to 0.64) when comparing performance of 20 medical students on two consecutive test sessions (Francis *et al.*, 2001).

These three methods of assessing dexterity have added objectivity in the assessment of surgical skill. The SAD was the first type of motion tracking tool, superseded by ICSAD which has the advantage of being a portable assessment device applicable to any surgical (or indeed non-surgical – a pilot study is underway with our Department to objectively assess skill when putting a golf ball) procedure. The ADEPT is bulky, cumbersome to use, and was not designed as an assessment tool for daily usage. The authors accept its use as an experimental and research tool.

Regardless of motion tracking devices, surgery is more than dexterity alone. It is necessary, and perhaps more important, to know whether the movements made are purposeful. For example, one may injure the common bile duct during a laparoscopic cholecystectomy – dexterity analysis alone would not record this potentially disastrous error. A surgeon may possess perfect technical skills, but to confirm competence it is necessary to analyse the context in which these movements are made. This leads to a discussion of video-based assessment in laparoscopic surgery.

1.5.3 Video-based assessment of laparoscopic skill

During the introduction of laparoscopic cholecystectomy, SAGES, SSAT and the EAES advocated proctoring of applicants for privileges in laparoscopic surgery by a qualified, unbiased staff surgeon experienced in general and laparoscopic surgery (Dent, 1991;Society of American Gastrointestinal Surgeons, 1991;European Association of Endoscopic Surgeons, 1994). As mentioned previously, evaluation of a trainee by the senior surgeon is open to subjectivity and bias, and hence unreliable (Elliot and Hickam, 1987;Van Rij et al., 1995;Warf et al., 1999). Adding criteria to the assessment can improve reliability and validity, typified by the Objective Structured Clinical Examination (OSCE) – a method to assess the clinical skills of history-taking, physical examination and patient-doctor communication (Cuschieri et al., 1979;Cohen et al., 1990).

1.5.3.1 The Objective Structured Assessment of Technical Skill (OSATS)

Martin et al. from the University of Toronto developed a similar approach to assessment of operative skill, the Objective Structured Assessment of Technical Skill, or OSATS (Martin et al., 1997). This involves six tasks on a bench format, with direct observation and assessment on a task-specific checklist, seven-item global rating score and a pass/fail judgement (figure 1i). 20 surgical residents of varying experience performed equivalent open surgical tasks on the bench format and live anaesthetised animals. Pearson's correlation between checklist and global scores for bench and live models were 0.81 and 0.87 respectively. Test-retest and inter-rater reliabilities were moderate for checklists (0.33 to 0.64) and higher for global scores (0.66 to 0.72). Global ratings were deemed to be more reliable and valid than checklists, consistent with previous studies performed by the same group (Winckel et al., 1994;Regehr et al., 1998).

Modifications of the global rating scale have been used by two studies to assess improvement in real laparoscopic performance when trained on either a box (Scott *et al.*, 2000a), or virtual reality simulator (Grantcharov *et al.*, 2004). However, studies to validate the scale have concentrated upon the assessment of open surgical skill, with none to specifically validate its use in laparoscopic surgery. The same group have though modified their scale for 27 senior general surgery residents to self-assess their performance following a laparoscopic Nissen fundoplication on live, anaesthetised pigs (Ward *et al.*, 2003). An operative component rating scale which is procedure-specific was also developed for this study, and demonstrated high inter-rater reliability (r=0.73 to 0.96) in the hands of expert surgeons.

Figure 1i - Detailed global 5-point rating scale and pass/failure score for Objective Structured Assessment of Technical Skills (OSATS)

Please rate the candidate's performance on the following scale:

Respect for tissue	1	2	3	4	5
	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments.		Careful handling of tissue but occasionally caused inadvertent damage.		Consistently handled tissues appropriately with minimal damage.
Time and motion	1	2	3	4	5
	Many unnecessary moves.		Efficient time/motion but some unnecessary moves.		Economy of movement and maximum efficiency.
Instrument handling	1	2	3	4	5
	Repeatedly makes tentative or awkward moves with instruments.		Competent use of instruments although occasionally appeared stiff or awkward.		Fluid moves with instruments and no awkwardness.
Knowledge of instruments	1	2	3	4	5
	Frequently asked for the wrong instrument or used an inappropriate instrument.		Knew the names of most instruments and used appropriate instrument for the task.		Obviously familiar with the instruments required and their names.
Use of assistants	1	2	3	4	5
	Consistently placed assistants poorly or failed to use assistants.		Good use of assistants most of the time.		Strategically used assistant to the best advantage at all times.
Flow of operation and forward planning	1	2	3	4	5
	Frequently stopped operating or needed to discuss next move.		Demonstrated ability for forward planning with steady progression of operative procedure.		Obviously planned course of operation with effortless flow from one move to the next.
Knowledge of specific procedure	1	2	3	4	5
	Deficient knowledge. Needed specific instruction at most operative steps.		Knew all important aspects of the operation.		Demonstrated familiarity with all aspects of the operation.

Overall, on this task, should this candidate: ☐Pass ☐ Fail ?

1.5.3.2 The Eubanks procedural rating scale

The advantage of an operation-component rating scale is that all parts of the procedure are assessed, whereas a global rating scale is generic and may ignore important steps of the operation. In this vein, Eubanks et al. developed a procedural rating scale for laparoscopic cholecystectomy, with scores weighted for completion of tasks and occurrence of errors (figures 1j and 1k) (Eubanks *et al.*, 1999). For example, liver injury with bleeding scores 5, whereas common bile duct injury scores 100. The error score is subtracted from the procedural score, and then divided by the highest procedural score possible; multiplying by 100 produces a percentage score. Three observers rated 30 laparoscopic cholecystectomies, performed by residents and attending surgeons. Correlation amongst observers for final scores was good ($r=0.74$ to 0.96, $p<0.05$), though correlation between final score and years of experience was only moderate ($r=0.50$, $p=0.057$).

Figure 1j – Score sheet used to calculate the raw scores for laparoscopic cholecystectomy

Initial Exposure	**Score**	**X If Completed**
Placement of fundus grasper	2	☐
Placement of body grasper	3	☐
Retraction of fundus cephalad	2	☐
Retraction of body anterolateral	3	☐
Initial Dissection		
Start dissection at body infundibular junction	5	☐
Identification of the cystic duct	5	☐
Circumferential dissection of duct	5	☐
Cystic Duct Dissection		
Adequate length of duct (enough for clips and catheter)	8	☐
Proximal clip	2	☐
Distal clip/ligature placement	2	☐
Division of duct	5	☐
Cystic Duct Cannulation (skip if cholangiogram not performed)		
Ductotomy	8	☐
Catheter placement	8	☐
Secure catheter	2	☐
Remove catheter	2	☐
Cystic Artery Dissection		
Identify cystic artery	5	☐
Circumferential dissection	5	☐
Adequate length (enough for clips and transection)	5	☐
Proximal clip	2	☐
Distal clip	2	☐
Transection of artery	5	☐
Gallbladder Fossa Dissection		
Areolar tissue division	10	☐
Inspect liver bed	4	☐
	Total Raw Points:	

Figure 1k – Error sheet used to calculate error points assessed during laparoscopic cholecystectomy. CBD, common bile duct.

Error	Error points	Frequency	Sum
Gallbladder			
Gallbladder Injury, Mechanical or Cautery (no bile spilled)	1		
Unintentional Release of the Gallbladder with Grasper	1		
Gallbladder Injury (bile or stones spilled)	10		
Liver			
Liver Injury (including cautery) without Bleeding	1		
Liver Injury with Bleeding	5		
Major Vascular Injury (other than cystic artery)	50		
CBD, Hepatic Duct Injury	100		
Cystic Duct			
Additional Attempt at Clip/Ligature Placement on Duct	1		
Additional Attempt at Ductotomy	1		
Additional Attempt at Cystic Duct Cannulation	1		
Misplaced Clip or Ligature on Cystic Duct	2		
Unintentional Removal of Cholangiogram Catheter	5		
Unintentional Cystic Duct Transection	10		
Failure to Cannulate Patent Cystic Duct (only mark once)	10		
Cystic Artery			
Additional Attempt at Clip Placement on Artery	1		
Additional Attempt at Cutting Cystic Artery	1		
Misplaced Clip on Cystic Artery (clip on clip, partial occlusion, and so forth)	2		
Mistaking Artery for Duct (or Duct for Artery)	5		
Cystic Artery Tear	15		
Miscellaneous			
Injury to Other Abdominal Viscus	25		
Prolonged Operative Time (>90 minutes, excluding cholangiogram)	10 pts/15 min		
		Total:	

1.5.3.3 Human reliability analysis

Joice et al. took a slightly different approach to laparoscopic procedural assessment (Joice *et al.*, 1998). They intended to document the nature and incidence of surgical errors enacted during laparoscopic surgery through utilisation of a modified Human Reliability Analysis (HRA) approach, which is the systematic assessment of human-machine systems and their potential to be affected by human error. This was based upon direct observation, and was adopted to categorise and record errors encountered during the practice of laparoscopic cholecystectomy. Eight surgical registrars were assessed undertaking a total of 20 laparoscopic cholecystectomies. The procedure was broken down into 10 steps (or tasks), such as 'dissect and expose cystic artery and cystic duct' and 'detach the gallbladder form liver bed'. Ten external error modes were defined and categorised into 'inter-step' (procedural) errors involved omission or re-arrangement of correctly undertaken steps, or 'intra-step' (execution) errors involved failure of a surgeon to correctly execute an individual step. The analysis revealed a total of 189 separate errors, of which 73 (39%) were inter-step and 116 (61%) were intra-step. The authors concluded that the high rate of inter-step errors were either due to a failing in the composition of task analysis, or a lack of adequate training to recognise the steps of the procedure and the order in which they should be applied. However, only 9% of the inter-step errors required corrective action, compared to 28% of the intra-step errors.

The study not only confirmed the applicability and usefulness of an observational methodology in the assessment of human error in endoscopic surgical performance, but the authors went on to identify aspects of the design and usage of instruments, surgical training and the differences between tasks which needed further directed research in order to identify underlying performance shaping factors (PSFs) and so reduce error rates. Further work from the same group has led to the application of an Observational Clinical Human Reliability Assessment (OCHRA) tool to laparoscopic cholecystectomy and pyloroplasty (Tang et al., 2004a;Tang et al., 2004b;Tang et al., 2005;Tang et al., 2006). Observational videotape data is subjected to a detailed step-by-step analysis of surgical operative errors, which are divided into consequential or inconsequential. This is a highly specialised and labour intensive task, though it has been suggested that OCHRA provides a comprehensive objective assessment of the quality of surgical operative performance by documentation of errors, stage of the operation when they are most frequent, and when they are consequential.

Rating scales are complex and time-consuming, the assessment of 20 surgical residents on OSATS requiring a total of 48 examiners for three hours each (Martin *et al.*, 1997). Furthermore, this method enlists direct observation of the subject, a factor that can lead to the introduction of bias and subjectivity (Elliot and Hickam, 1987;Van Rij et al., 1995;Warf et al., 1999). Video-based analysis using any of the tools described can reduce this bias, but still requires the time and concentration of expert surgeons to assess the procedure. The aim is to develop a feasible, objective, reliable and instantly accessible review of performance relating to surgical skill. Virtual reality simulator devices have been suggested as a possible tool for this purpose.

1.5.4 Virtual reality simulators for laparoscopic skills assessment

Studies to confirm the role of virtual reality simulators as assessment devices have concentrated upon the demonstration of construct validity, i.e. a test of whether the model measures what it purports to measure, and is extrapolated from its ability to differentiate between different levels of experience.

1.5.4.1 Minimally Invasive Surgical Trainer – Virtual Reality (MIST-VR)

In 1998, Taffinder et al. reported the first study to validate the scoring system of MIST-VR according to different levels of surgical experience (Taffinder *et al.*, 1998). Experienced surgeons (>100 laparoscopic cholecystectomies) were found to be significantly more efficient, made less correctional sub-movements and completed the virtual reality tasks faster than trainee surgeons or non-surgeons. Following on from this study, McNatt and Smith tested three experienced (>250 cases) and five novice laparoscopists on the MIST-VR platform (McNatt and Smith, 2001). Each group performed three tasks (target acquisition, target traversal and target manipulation with diathermy) with a minimum of seven repetitions. Within each task, time, errors, and economy of movement for each hand were assessed. The experienced group were significantly quicker, made fewer errors and had greater economy of movement than the novices on tasks one and three, but only excelled in terms of fewer errors on the second task. Nonetheless, it was apparent that this device could discern levels of laparoscopic manipulative skill according to experience.

Gallagher et al. in 2001 recruited 12 experienced (>50 laparoscopic procedures), 12 inexperienced (<10) and 12 novice laparoscopic surgeons to perform one run of six tasks on the MIST-VR simulator (Gallagher *et al.*, 2001). The experienced group completed the tasks significantly faster, had a lower error rate and were more economic in their movement of surgical instruments and in their use of diathermy($p<0.01$). As well performing to a higher standard, the experienced surgeons as a group were also more consistent in their performance – an important point when considering surgical competence. Gallagher confirmed these findings with a further study published in 2002 (Gallagher and Satava, 2002), as did Grantcharov et al. in 2003 (Grantcharov *et al.*, 2003). As previously mentioned, these last two studies also confirmed the fact that the novice group required the greatest numbers of sessions to reach steady-state. More recently, the performance scores of experienced surgeons have been used as benchmark values, the intention being to provide a marker for trainees to achieve, and thus attempt to define proficiency levels prior to progression onto real cases (Gallagher *et al.*, 2004).

Further evidence for the validity of MIST-VR has hailed from the confirmation of concurrent validity, whereby 32 first- and second-year medical students were taken into a porcine laboratory in order to assess two operative tasks (measuring a piece of bowel and placing a piece of bowel into a laparoscopic bag) (Madan *et al.*, 2005). Then they were taken into an inanimate lab with a MIST-VR and each student repeatedly performed one task (placing a virtual reality ball into a receptacle). The students' scores and times from the animate lab were compared with average economy of movement and times from the MIST-VR, and there was statistically significant ($p<0.05$) correlation between 11 of 16 possible relationships between the virtual reality trainer and operative tasks. The authors of this study suggest that 'virtual reality may be an avenue for measuring laparoscopic surgical ability'.

At the American College of Surgeons meeting in 2001, Gallagher et al. investigated the performance of 210 experienced laparoscopic surgeons on two trials of MIST-VR (Gallagher *et al.*, 2003b). The aim was to benchmark performance of these surgeons, to confirm future use as an assessment tool. The results revealed marked variability in the scores obtained, together with a significant learning effect between trials. In order to use such data for high-stakes assessments, perhaps a pool of expert scores from all centres currently using VR simulation can lead to the development of an international benchmark for trainee surgeons. Furthermore, as Seymour et al. have commented upon, some trainees take longer to achieve pre-defined levels of proficiency than others (Seymour *et al.*, 2002). This can enable particularly gifted trainees to be fast-tracked into an advanced laparoscopic program, and the true development of a competency rather than time-based curriculum.

1.5.4.2 LapSim

Similarly, studies to develop the use of this simulator as an assessment device have focused upon construct validity. The first article to be published was by Hassan et al., in German, which recruited 27 physicians from three experience groups (experienced, intermediate and novice) (Hassan *et al.*, 2005). Each participant performed seven basic laparoscopic skills tasks twice, with the results confirming superiority of the experienced group in 'most' of the tasks. An important comment from this study was that the second run revealed greater differences between the groups than the first run, since there is a period of familiarisation to the simulator, even for the experienced surgeons. Furthermore, tasks of low complexity in virtual reality such as camera navigation did not show significantly different results between the three groups.

Duffy et al. tested the performance of 54 subjects on eight tasks (seven basic skills plus intracorporeal suturing) on the simulator (Duffy *et al.*, 2005). Not unreasonably, the greatest group differences were between experienced and novice surgeons, and the task which exhibited greatest differences was intracorporeal suturing – the most complex. More recent studies published by Langelotz et al. (Langelotz *et al.*, 2005), Eriksen et al. (Eriksen and Grantcharov, 2005) and Woodrum et al. (Woodrum *et al.*, 2006) have corroborated the construct validity of the LapSim device as a tool for assessment of laparoscopic surgical skill. Contrary to this evidence, Ro et al. found novices to have better performance compared to the experienced basic laparoscopists during their first exposure to the LapSim basic skill and Calot's triangle dissection modules (Ro *et al.*, 2005). All subjects were provided with an initial practice session which would discount the suggestion that the result is due to a lack of familiarisation with the tool.

1.5.4.3. Other laparoscopic virtual reality simulation devices

There is one published study regarding the construct validity of the basic skills tasks on the LapMentor virtual reality simulator device (McDougall *et al.*, 2006). McDougall et al. recruited 103 subjects divided into medical students (n=23), residents/fellows (n=24), and experienced laparoscopic surgeons performing either <30 cases (n=26) or >30 cases per year (n=30). Once again, assessment parameters were able to differentiate through the experience levels, though there were only significant differences between all groups of subjects for the most complex task. Of the other virtual reality simulator devices, construct validity studies have been published for ProMIS (Van Sickle et al., 2005;Broe et al., 2006) and Simendo (Verdaasdonk *et al.*, 2006).

A fundamental issue with all of these construct validity studies is the use of a variety of assessment parameters, statistical analysis and group definition. This makes it extremely difficult to perform a systematic analysis of the aforementioned studies, which is the ideal manner of defining levels of evidence. In response to this, the EAES Work Group for Evaluation and Implementation of Simulators and Skills Training Programmes have published 'Consensus guidelines for validation of virtual reality surgical simulators' (Carter *et al.*, 2005). Regular updates will be necessary, but this is a concerted effort to inform the surgical community of the use of these devices.

In essence, the use of virtual reality simulation to assess laparoscopic skill can enable surgeons to be assessed on a standard procedure, without encountering the difficulties of patient variability and disease severity. This can be carried out at any time and does not require further processing or manpower to produce a test score. The tasks are standardised, enabling comparison of results with those of other surgeons and institutions across the globe, and recorded, to highlight areas requiring closer attention.

1.5.5 Evaluation of assessment tools

There is a consensus agreement to strive toward objective assessment of surgical skill (Satava, 1999). Analysis of dexterity, procedure-based and video-based rating systems are all promising in their own right. However, there is no consensus as to the optimal method of assessment for laparoscopic procedures, and perhaps there is a role for all modes to be used in conjunction, possibly in a hierarchical manner. The role of virtual reality devices as assessment tools is likely to be viable for basic skills performance, but higher level assessments probably require more complex models and definitive assessment parameters.

In the vein of developing new tools, our department has recently developed new software to enable the ICSAD trace to be viewed in unison with a video of the procedure, leading to a dexterity-based video analysis system. This still requires an investment of time to assess the procedure on a rating scale, but it may be possible to use motion analysis to identify areas of substandard performance, and concentrate upon assessing these areas alone. Ward et al. enabled examiners to fast forward through videotapes at their discretion, only assessing parts of the procedure they deemed important, without sacrificing reliability (Ward *et al.*, 2003). In order to maintain standardisation of the assessment though, it may be possible to use other tools such as dexterity analysis to highlight areas of weak performance. For example, if one part of the operation had a greater number of movements than other parts, the surgeon may have been experiencing difficulties, leading to closer scrutiny of this part of the video. Conversely, remaining parts of the video may not require such extensive analysis. Other tools currently in development may also be used to highlight important areas of the operation, such as eye-tracking devices (Dempere-Marco et al., 2006).

1.6. Conclusions

'No surgical technique in recent memory has generated as much excitement and enthusiasm among general surgeons as has interventional laparoscopy...the access may be minimal, but the operations and potential for complications are major' (Dent *et al.*, 1991).

The introduction of laparoscopic surgery in the past 15 years has had a significant impact on modern surgery. Laparoscopic cholecystectomy is still the most common procedure, but the list of laparoscopic advances is steadily growing. With these advances comes the need to assess competence of the surgeon carrying out the procedure. This does develop with experience, but the days of accomplishing the learning curve at the expense of patients' welfare are no longer appropriate (Reznick and MacRae, 2006).

In order to ensure that virtual reality simulation can be incorporated into current training programmes, there is a need to develop validated curricula for basic, intermediate and advanced level laparoscopic training. Current simulators enable basic and some forms of intermediate level training, though the methods to achieve this remain unclear. We must know how often training should occur, and whether tutors are required to be present at all times. Competency levels should be defined using at least a national approach, enabling standardisation of training programmes.

It is also important to ensure that virtual reality simulation is seen as an adjunct to traditional methods of training, and not as an alternative. Virtual reality simulators currently have the ability to teach basic laparoscopic skills, enabling novice surgeons to progress along the early part of the learning curve prior to entering the operating theatre. With further developments in virtual reality simulation technology, it may be possible to practice complete procedures such as Nissen fundoplication and colectomy. However, surgeons will still need to reach expert levels of skill in the operating theatre, and further training to achieve this is essential.

The organisation of training programmes must also be addressed. It is not necessary for every hospital have a virtual reality simulation laboratory, rather for training to be provided at regional skills centres. Currently, training occurs in the form of isolated two or three-day courses, but with the establishment of regional centres, it would be possible to develop training programmes over a period of months and years. Junior surgeons would initially attend an intensive basic laparoscopic skills programme, followed perhaps by one session per week to learn intermediate and advanced skills. This can be aligned with their operating theatre experience, and their performance charted using the previously mentioned assessment tools.

To confirm the present situation, laparoscopic simulators must be assessed in a competency-based manner, and not only show improvement in performance at the first laparoscopic cholecystectomy, but whether this effect persists and for how long, the ultimate aim being to prove a shortening of the learning curve in the operating theatre to achieve surgical proficiency in the same way that pilots do on the ground prior to flying their aeroplanes. Once irrefutable evidence has been compiled for the use of laparoscopic simulators in surgery, the next steps could be the use of simulators for teaching advanced laparoscopic procedures, and for training in crisis situations. Then, not only would they be useful for inexperienced surgeons, but the surgical community as a whole.

In 2004, an Editorial commented on surgical simulation as a 'good idea whose time has come' (Champion and Gallagher, 2003). Professional bodies too are excited about the future of virtual reality simulators. The American College of Surgeons states that 'simulations would reduce error by...screening of potential surgeons for demonstrable aptitude, provision of initial training in surgical experience, promotion of the ongoing education of surgeons, and maintenance of proficiency'. Though surgical simulators are growing out of the research phase, there is a growing confidence in their ability to revolutionise surgical education in the future.

In the meantime, it is important to ensure the lessons learnt from the introduction of laparoscopic surgery are not repeated. The widespread introduction of laparoscopic procedures should be in a controlled and audited manner. Training programs should exist and proctoring must remain an important part of the learning process. Only with this structured and defined approach can we hope to avoid the complications of the last decade.

Chapter 2

A FRAMEWORK FOR SYSTEMATIC TRAINING AND ASSESSMENT OF TECHNICAL SKILL – STATS

Aim

Develop a standardized and universal methodology for development of curricula for the acquisition of technical surgical skills, based upon motor learning theories, task analysis and transfer of skill to the operative environment.

2.1 Introduction

The purpose of training programmes for all medical specialities are to produce individuals who are competent to meet the healthcare needs of society. Recent Editorial publications have commented upon the crisis in medical education, and the requirement for defined competencies to assess performance prior to independent medical practice (Silen, 2001;Weigelt, 2003). Effective since July 2002, The Accreditation Council for Graduate Medical Education (ACGME) listed six categories of competence, defined as the ACGME Outcomes Project, i.e. patient care, medical knowledge, practice-based learning, interpersonal and communication skills, professionalism and systems-based practice (Pellegrini, 2002).

Though positive, there is not a prescription for how to train or indeed assess skill in each of these categories. Specifically, the need for demonstration of proficiency in technical skills is not addressed. This is not only an issue for the surgical specialities, but also for physicians training in cardiology, anaesthesiology, gastrointestinal medicine, chest medicine, and interventional radiology, together with allied health specialists. The introduction of new techniques and instruments to the medical world requires training of not just residents, but also of certified practitioners. This was clearly evident with the increased rate of complications associated with the introduction of laparoscopic cholecystectomy, and has led to the development of training programs at many centres around the globe (1991;Powers *et al.*, 2002).

Table 2a – Core competencies for the ACGME ‘Outcomes Project’

Patient Care
Medical Knowledge
Practice-based learning
Interpersonal and communication skills
Professionalism
Systems-based practice

Within the arena of laparoscopic skills training, practice on simulation-based devices such as box-trainers and virtual reality systems have been shown to lead to improved performance within the clinical domain. However, each study used a different training curriculum with varied end-points. The majority of studies requested subjects to perform the tasks a pre-set number of times (Scott et al., 2000a;Hamilton et al., 2002;Grantcharov et al., 2004). Seymour et al. set benchmark criteria on one of the MIST-VR tasks and subjects had to achieve these criteria prior to performing the test laparoscopic cholecystectomy (Seymour *et al.*, 2002). This was the first time that subjects had to demonstrate their skill, ensuring that it was not number of hours or sessions spent training which defined proficiency, but demonstration of a pre-set level of skill.

Though a commendable application of the simulator device, current training schedules have not been designed from a background of scientific research to ensure the curriculum is valid, efficient and competency-based. The aim of a surgical residency program is to produce competent professionals, displaying the cognitive, technical, and personal skills required to meet the needs of society. Within the context of surgical procedures, patients are at liberty to expect satisfactory outcomes in terms of cure, complication rates and return to daily activities. Technical proficiency is paramount to the delivery of such outcomes, and sub-standard performance must be recognised, enabling review and possible modification of the training program.

Laparoscopic simulation companies are only just beginning to produce whole-procedure software, thus it is necessary to identify the key skills, and tasks, which must be learnt prior to performing a whole procedure. This is the basis of a hierarchical task-based approach to acquiring surgical skills (Cao *et al.*, 1999). The aim is to deconstruct a procedure into its constituent parts, and develop a systematic and step-wise strategy for learning the individual skills required. However, surgery is a cumulative process, and thus within such a framework, it is possible to identify key tasks which are integral to the procedure. These tasks may also be applicable to other procedures, e.g. during laparoscopic appendicectomy, many of the same skills are used for dissection of the mesoappendix to identify the appendicular artery as in dissection of Calot's triangle for identification of the cystic duct and artery during laparoscopic cholecystectomy. Identification of key tasks may provide a suitable tool for assessment which can be feasible, valid and reliable (Annett et al., 2000;Payandeh et al., 2002). The aim was thus to develop a stepwise, systematic and competency-based strategy for both training and assessment of technical skills, both within the skills laboratory and transferring to the clinical environment (Kneebone *et al.*, 2004).

2.2 Methods

A literature search was performed on PubMed and Medline from its inception to present day for all English language articles using the following MeSH headings: education, professional; curriculum; clinical competence; teaching, and keywords of skills, assessment and simulation. The following educational and nursing databases were also searched for relevant articles: the Education Resources Information Centre (www.eric.ed.gov) database, the Cumulative Index to Nursing and Allied Health Literature (CINAHL) and the British Nursing Database (BNI), using the same MeSH headings as for the initial search.

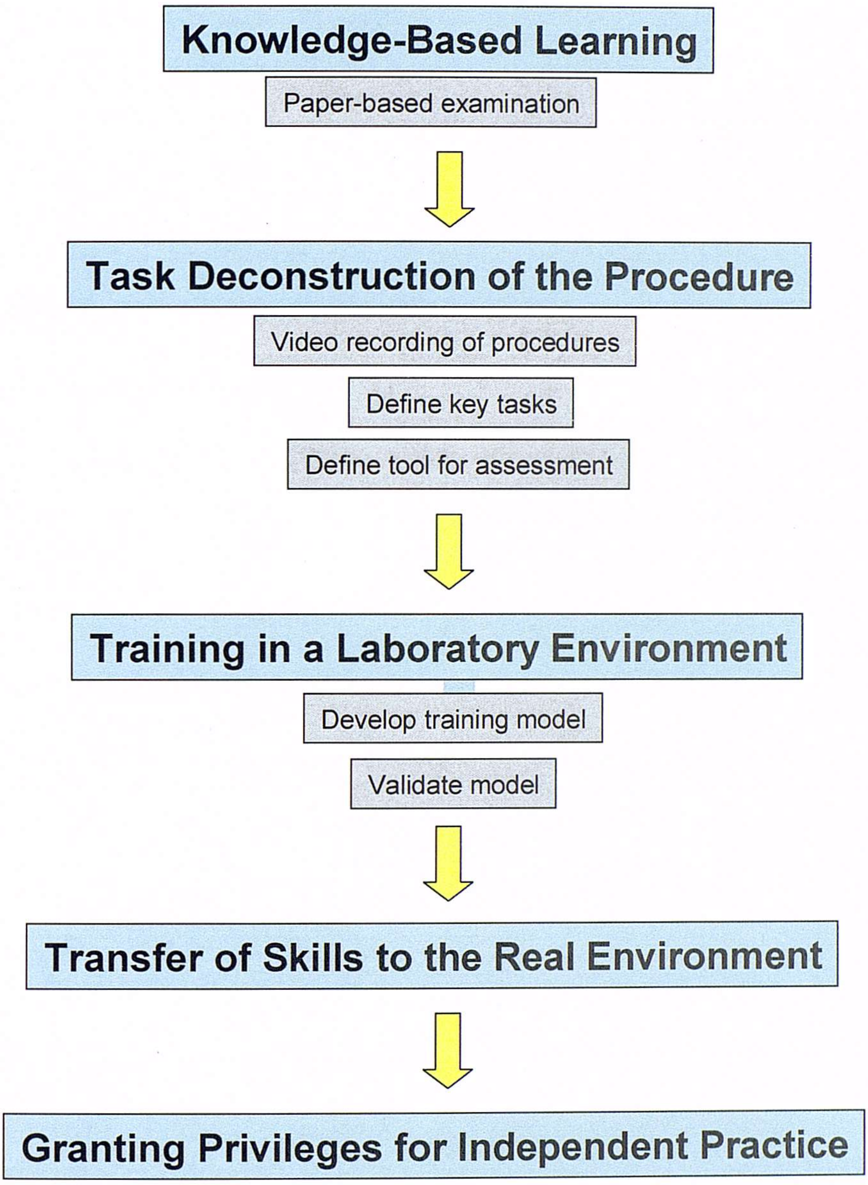
Further articles were obtained from references within papers identified by the initial search, in addition to articles from the authors' experience and by discussion with other experts in the field of medical education.

2.3 Systematic training and assessment of technical skills (STATS)

The outcome of the literature search enabled the definition of a strategy for surgical technical skills training and assessment. The framework was divided into four sections, each of which is described.

1. Knowledge-based learning
2. Task deconstruction of the procedure
3. Training in a laboratory environment
4. Transfer of skills to the real environment

Figure 2a – A Framework for Systematic Training and Assessment of Technical Skills (STATS)



2.3.1 Knowledge-Based Learning

A systematic approach to learning must begin with the acquisition of procedure-specific knowledge. Though common sense, Hassan et al. investigated this proposition by training two groups of novice laparoscopists: medical students with neither knowledge nor technical skill, and junior residents with laparoscopic knowledge, though no practical experience (Hassan *et al.*, 2006). Pre-training scores on the 'cutting' task of the LapSim simulator were similar for both groups, though after training the junior residents were significantly faster, more precise and caused reduced trauma to the tissues than the medical students. The authors conclude that 'clinical background and understanding of the clinical value of a training program lead to faster acquisition and improvement of laparoscopic skills as performed on the laparoscopy simulator.' None of the studies to assess the use of simulation-based device to acquire laparoscopic skills have considered cognitive knowledge as part of the training curriculum. Instead, subjects are 'familiarised' to the simulator and it is only technical skills which are taught and assessed.

The delivery of knowledge-based learning for the acquisition of technical skills can be broadly (and arbitrarily) defined into the categories of:

1. pre-procedure assessment and preparation,
2. anatomical knowledge,
3. safety and limitations of specific instruments,
4. ergonomics, and
5. post-procedure management.

The knowledge requirements for each procedure (or procedure type, e.g. laparoscopic) for these categories can be defined by a consensus group of experienced surgeons. This forms the basis of a syllabus, one example of which is the Surgeons in Training Education Module (or STEP), published by The Royal College of Surgeons of England as a distance-learning programme (2006). These must be studied by the trainee prior to technical skills training, followed by satisfactory performance on a paper-based or electronic test of knowledge which is necessary to ensure that the required information has been learnt, for example, those required for membership to any of the Royal Colleges in the United Kingdom.

Within this mode of learning, it is crucial to teach likely errors which can occur during an operative procedure, enabling trainees to anticipate and avoid errors, or to identify when an error has been committed, together with strategies to manage them (Rogers *et al.*, 2002).

However, the reality of how one learns is far removed from the simplistic notion of acquiring procedure-specific knowledge. Learning and educational theories have been proposed in an attempt to make students learn more effectively (Jolly and Rees, 1998). For example, Ausubel et al. stated that the 'most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly' (Ausubel *et al.*, 1978). There are few who would argue with this concept, and indeed a structured, hierarchical approach to learning which builds upon past knowledge is the basis of the apprenticeship model of training, popularised by Halstead at the end of the 19th century (Halsted, 1904).

Kolb defined the experiential learning cycle, enabling a learner to reflect upon their own experience, develop a theoretical understanding of it, which can be applied to new situations, providing further opportunities for experience and reflection (Kolb, 1984). This model appreciates the intuitive nature of medical practice, and the importance of defining otherwise hidden theories of action. It is though well known in medicine, especially during technical tasks, that the underlying mechanisms of performing a skilled procedure are a mystery even to the experienced operator (Patkin and Isabel, 1995). For an experienced surgeon, tying a surgical knot is 'easier *done* than said'. To teach an inexperienced trainee to perform a procedure necessitates a rigorous mechanism for elucidation of the steps of the procedure, which are then taught in a structured manner. When the student has acquired the basic ingredients of a technical specialty, experiential learning should enable the application of this knowledge to new situations. The teacher can be there to facilitate and guide the learning process, in a supportive manner.

This model is echoed by Fitts and Posner who described three stages of motor skills learning (Fitts AM and Posner MI, 1967). During the first stage (cognition), the trainee gains an understanding of the task through instructor explanation and demonstration. This relates to the quality of the human-machine interaction, resulting in a short phase for a realistic model. In the second stage (association), the user practices the task to eliminate error, with the instructor providing feedback to identify errors and suggest corrective actions. The final phase (automation) occurs when the learner performs the tasks in a relatively automated fashion with little or no cognitive input.

In order to achieve improvements in technical skill, the first or cognitive phase must be transgressed, which relies upon procedural knowledge. If the subject understands the purpose of the task, then they should be able to learn the skills in an efficient manner. A lack of understanding simply leads to a machine-based approach to learning, during which the learner may lose motivation. Within this frame, is to ensure contextualisation of the task in hand. Knowledge-based learning is focused upon the real task, and if the learner is then subjected to an abstract task (such as on MIST-VR), it may be difficult to understand the relevance of the practical session. For basic skills training, an important though declining issue is of poor motivation to learn on bench-top models (Grober et al., 2004a). A recent survey of general surgery program directors in fact reported 88% to consider skills laboratories an effective mode of improving operating room performance (Korndorffer, Jr. *et al.*, 2006). For example, during laparoscopic skills training, it may be difficult for the learner to grasp the clinical relevance of transferring chick peas from one pot to another. It may be necessary to begin in a concrete manner by providing trainees with relevant first hand experiences of surgical practice, prior to the presentation of conceptual training tasks (Mayer, 1979).

Training to the automated phase with associative experiential learning leads to competence, but rarely to expertise. Ericsson describes the notion of *deliberate* *practice* as a critical process for the development and maintenance of mastery or expertise (Ericsson, 2004). The focus is upon selection of a defined task to improve overall performance through repeated practice and supportive feedback. Time devoted to deliberate practice has been shown to relate to attained levels of expertise in other high performance domains, i.e. expert chess players, musicians and athletes. Within a surgical context, operative procedures are complex, and it is much more manageable to focus on a specific part of the procedure for further training. Furthermore, Ericsson emphasises the role of life-long learning to ensure maintenance of expertise.

2.3.2 Task Deconstruction of the Procedure

In order to develop a systematic approach to technical skills training, trainees must learn skills in manageable, bite-size pieces. This can then form a hierarchical process to skills acquisition, in a not dissimilar manner to the Halstedian apprenticeship model (Halsted, 1904). However, this model of training was determined by the senior surgeon, thus was not standardised and makes it difficult for trainees at commencement of their surgical career to acquire skills. Trainees need to learn a standardised method to complete a task or procedure, beyond which modifications can be made according to specific nuances which they may have acquired or discovered through the learning process.

2.3.2.1 A description of task analysis

A task analysis of any procedure can be performed through review of videotaped procedures, training manuals, technical protocols of the operation and discussion with experienced operators, initially described by Joice et al. for laparoscopic cholecystectomy (Joice *et al.*, 1998). Sarker et al. performed a similar analysis, through use of textbooks, articles, papers, web pages, surgical skill course manuals, and expert panel discussions during which the analysis could be evaluated and modified if required (Sarker *et al.*, 2006). This point is important, as different strategies are regularly employed by different operators and it is thus necessary to observe procedures performed by a number of experienced clinicians. Laparoscopic surgery makes this quite amenable by observation of videotaped procedures, followed by division into steps and sub-steps. As well as task decomposition of laparoscopic cholecystectomy, MacKenzie et al. have described an extensive analysis for the procedure of Nissen fundoplication (MacKenzie *et al.*, 2001).

In terms of a hierarchical approach, the procedure can be divided into tasks, followed by steps and sub-steps if so desired. The level of complexity depends upon the desired output of the process. For example, a task-based approach may be adequate for definition of a training programme for junior surgeons to learn a surgical procedure, though identification of errors during a procedure will require a more detailed analysis of each step and possibly sub-step of the procedure. MacKenzie et al. fragmented the procedure of Nissen fundoplication into seven tasks, as shown below in table 2b (MacKenzie *et al.*, 2001).

In order to standardise the beginning and end of each task for operator and observe alike, clear and unambiguous start and end points must be defined. For example, step 6 of 'Wrap fundus' begins with 'moment oesophageal elevator contacts abdomen' and ends at 'completion of last suture – endostitch is removed'. MacKenzie et al. performed a more detailed task analysis, down to the level of 'position jaws', 'bite tissue', 'pull needle through' and so on (MacKenzie *et al.*, 2001). The authors state in an earlier publication that 'the hierarchical decomposition of surgical procedures provides a framework for structuring a systematic approach to training, in the real and simulated environment' (Cao *et al.*, 1999).

Table 2b – Operational definitions of surgical steps in Nissen fundoplication

Surgical Steps	Beginning	Ending
1. Prepare patient	Moment the insufflation needle contacts abdomen	Liver in place and liver elevator is stable
2. Divide peritoneum	Moment the tool moves towards the peritoneum to be cut	Last cut of peritoneum
3. Expose crura and GE junction	Last cut of peritoneum	Last cut of tissue and when scissors are removed
4. Repair crura	Moment the endostitch contacts abdomen	Completion of cut suture and removal of endostitch
5. Divide short gastrics	Moment scalpel contacts abdomen	Last cut and scalpel removed
6. Wrap fundus	Moment oesophageal elevator contacts abdomen	Completion of last suture – endostitch is removed
7. Close	When endostitch is removed	Completion of last open suture

2.3.2.2 Application of task analysis to a training curriculum

The division of a procedure into constituent parts can have a number of applications in training and assessment of surgical technical skills, examples of which are offered in the text below:

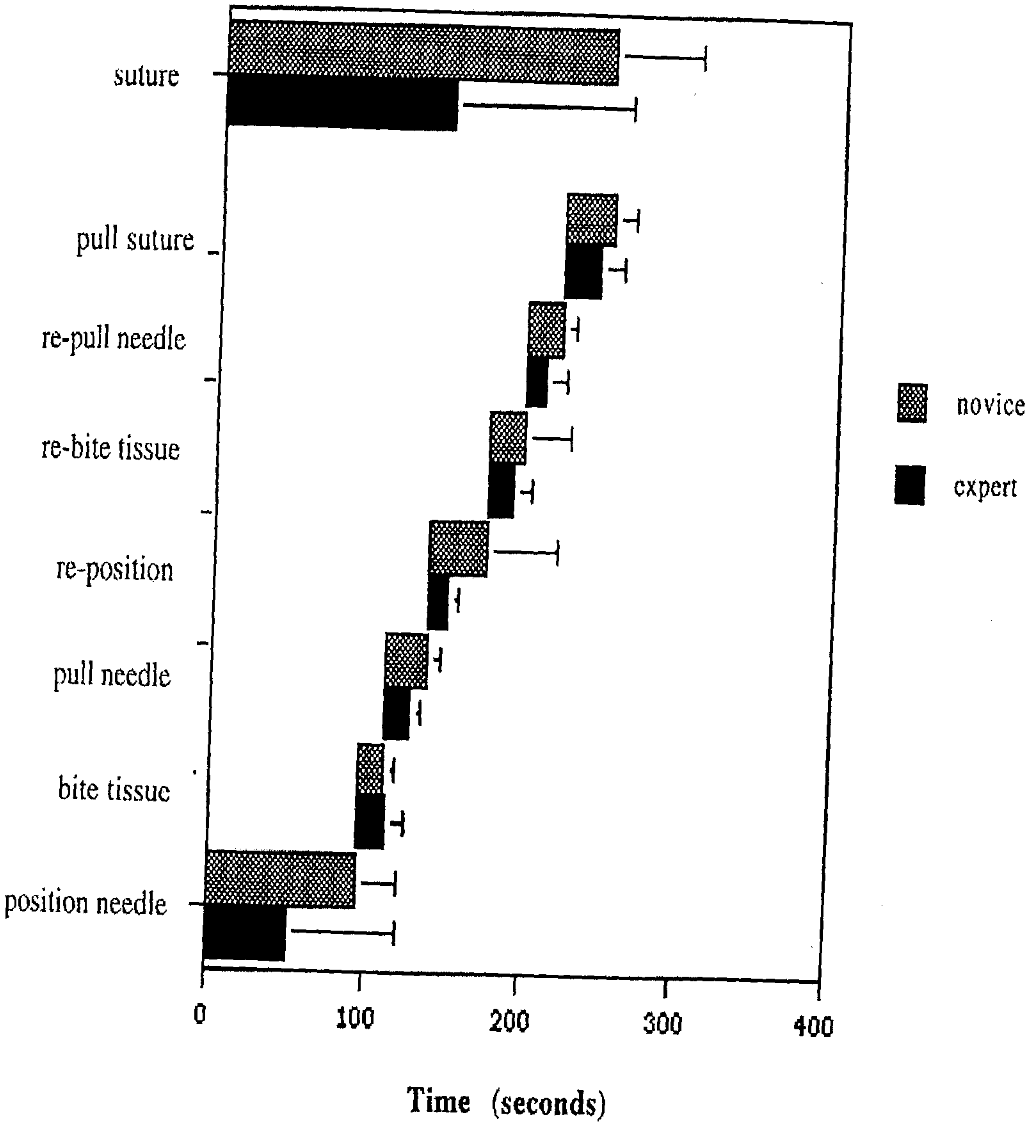
a) Identify tasks common to different procedures
b) Identify procedure-specific tasks
c) Define key tasks of a procedure or procedure-type
d) Define a hierarchical structure to task-based training
e) Direction of resources for simulation of key surgical tasks
f) Error analysis

2.3.2.2.1 Identify generic surgical tasks: 'Clip and cut' is a task which has been utilised in a number of box-trainer and virtual reality training studies, and is generic to a number of laparoscopic procedures from cholecystectomy to appendicectomy through to colonic resections. Similarly, dissection of the oesophageal hiatus during a Nissen fundoplication utilises similar manoeuvres to dissection of the rectum from the pelvic side walls. Though the anatomical considerations are different, similar skills are necessary to achieve the desired effect.

2.3.2.2.2 Identify procedure-specific tasks: In contrary to identification of generic tasks, there may be some tasks which are procedure or organ-specific, e.g. performing an intra-operative cholangiogram. Though the skills required to perform the task remain generic, specific steps or tools must be used to acquire proficiency.

2.3.2.2.3 Define key tasks of a procedure or procedure-type: Through comparative observation of inexperienced and experienced operators performing the same procedure, it is possible to define the most challenging parts of the procedure. Payandeh et al. adopted a task-based approach to assess performance at intracorporeal suturing of expert and novice laparoscopic surgeons on a bench-top model (Payandeh *et al.*, 2002). The task was divided into seven steps and the time taken for each subject to complete the step recorded. Compared with expert surgeons, novices took longer, on average, to complete all the subtasks. A statistical analysis was not presented, though the chart of the timelines reveals greater differences between novices and experts in two of the seven tasks – namely position and re-position of the needle (figure 2b).

Figure 2b – Timeline for the suturing task, and the subtask components.



Though an inference, it is suggested that these are the two most challenging steps to master, and acquisition of technical skill to expert levels in these two steps would perhaps enable a trainee to perform the entire task (from a technical point of view) in a time similar to that of the expert surgeon.

Though critics would suggest that time is not an appropriate measure of surgical performance, the methodology of this study is sound and makes it possible to define the most challenging tasks of any procedure (Payandeh *et al.*, 2002). Researchers should aim to use objective assessment tools, for example dexterity measures and rating scales, in order to provide further information regarding the differential performance of experienced and novice operators (Darzi *et al.*, 1999). Expertise in the key tasks can then be compared to overall procedural performance, and correlations calculated using the same assessment measures for both task and procedural performance.

In a paper by Beard et al., a vascular procedure commonly performed by trainee surgeons was deconstructed into five simulated tasks (Beard *et al.*, 2005). Thirty-three general surgical trainees undertook the five simple simulations (knotting, skin incision and suturing, tissue dissection, vessel ligation and small bowel anastomosis). The operative competence of each trainee was then assessed during two or three sapheno-femoral disconnection procedures performed on patients, by a single surgeon on a rating scale. The best predictors of operative competence were reported to be the number of sapheno-femoral disconnection procedures performed previously, and the simulation scores for dissection and ligation tasks. The authors state that 'deconstruction of operations into their component parts enables trainees to practise on simple simulations representing each component, and be assessed as competent, before undertaking the actual operation'. This can enable both surgical educators and simulation developers to build models of key tasks, not only for training but also for assessment of surgical proficiency, obviating the current zeal for development of whole-procedure simulations.

2.3.2.2.4 Define a hierarchical structure to task-based training: Task deconstruction and subsequent definition of task difficulty can lead to a step-wise and graded mode of training whereby easier tasks are performed at commencement of the curriculum, leading onto more challenging tasks. This is an evidence-based approach to structured training curricula, and can be actioned by comparing the learning curves of novice and experienced subjects on a surgical task. For example, McDougall et al. recently observed that higher level skills tasks on the LapMentor simulator were better at differentiating study groups than some of the simpler tasks (McDougall et al., 2006). Within a laparoscopic skills training curriculum, novice subjects should begin with the simpler tasks, moving onto the complex tasks when deemed appropriate, preferably through objective assessment of skill. However, such an evidence-based laparoscopic training curriculum does not yet exist.

The acquisition of basic skills such as suturing and knot tying is fundamental prior to progression onto more complex tasks. In the surgical skills laboratory, these tasks are taught in a context-free manner, though in real situations, procedure-specific knowledge (and practice) may be necessary to enable the trainee to modify the task accordingly. This underlines both the experiential model of learning, and the notion of deliberate practice, as previously described (Kolb, 1984;Ericsson, 2004). The aim is not to try to produce expert surgeons solely through simulation-based training, but rather to arm trainees with a basic semblance of surgical skill prior to entering the operating theatre, subsequently shortening the time taken to achieve satisfactory outcomes within the clinical environment.

2.3.2.2.5 Direction of resources for simulation of key surgical tasks: Task analysis may provide simulator developers with information regarding how to direct their resources (Seymour and Rotnes, 2006). The development of tasks common to a number of different procedures is a commercially attractive proposition, though must be balanced with the need for anatomically specific (or face valid) simulations.

2.3.2.2.6 Error analysis: Joice et al. adopted this approach in order to define a skills assessment methodology for laparoscopic cholecystectomy (Joice *et al.*, 1998). Fragmentation of a procedure into its constituent parts enables assessment of error, though is also useful when developing a training model to define errors which may occur.

2.3.3 Training in a Laboratory Environment

Definition of tasks of a procedure or procedure types must lead to structured and standardised training, commencing in a skills laboratory environment. The ideal platform for training may be said to be the patient itself, though this can be problematic for learning new skills due to patient and disease variability, and is considered unethical (Reznick and MacRae, 2006). The development of technical skills laboratories and advancements in simulation technology can allow trainees to learn the skills required in a safe environment on standardised models, which should then transfer to improved performance in the real environment (Scott et al., 2000a;Seymour et al., 2002;Sedlack and Kolars, 2004;Grantcharov et al., 2004).

2.3.3.1 Models for Laboratory-Based Training

The models to be used may be of synthetic, cadaveric or live animal tissue, or computer-based. The choice of model-type is defined by cost and availability. Synthetic models are relatively cheap, easy to use and store and can teach the skills required for mastery of a wide variety of technical skills such as suturing, central venous catheter placement and episiotomy repair. Synthetic models also exist for training of simple surgical procedures such as hernia repair, appendicectomy, cholecystectomy and Nissen fundoplication. However, they have been criticised for a lack of realism of tissues, and problems with contextualisation of the full procedure (Kneebone *et al.*, 2004).

Animal tissue can provide a greater sense of realism for training in technical skills, though necessitates specialist storage and handling facilities (Clayden, 1994). Cadaveric animal tissue has greater fidelity than synthetic models, and can allow trainees to learn the skills required in a classroom environment. Anaesthetised animals can also be used to teach the skills required to perform a complete procedure, though this form of training is illegal in the United Kingdom (Byrne, 1994). The simulation can also cause problems due to differences in anatomical configurations of the animal – an example is the lack of a mesenteric attachment of the sigmoid colon when practicing a colonic procedure on a porcine model.

The field of training using virtual reality (VR) simulation currently encompasses systems to teach laparoscopic, endoscopic and percutaneous interventional techniques (Krummel, 1998;Satava, 2001b). The models are standardised, and following the initial outlay cost are relatively cheap to maintain. The further advantage of VR simulation is the in-built and automated feature of objective assessment, providing data on parameters such as number of errors, instrument path length and quality of the procedure performed. During a simulated coronary catheterisation, it is possible to provide data on the extent of the lesion traversed with the stent, or for colonoscopy, percentage of red-out time, i.e. the time that the endoscope tip is pressed against the wall of the colon and thus not performing a useful function.

Decisions regarding which simulation modality to use depend upon the intended outcomes of the training curriculum, as previously described. The fidelity of synthetic, animal and VR models is also variable, and the general consensus has been to use the highest fidelity possible. However, Grober et al. have called into question the need to train on the highest fidelity models available, with equivalence of learning on synthetic and anaesthetised rat models for spermatic cord micro-anastomoses (Grober et al., 2004b).

2.3.3.2 Validation of simulated models

It is both time-consuming and expensive to simulate an entire procedure, and indeed for the acquisition of technical skills it may not be necessary to do this to achieve the desired goals. However, the notion of developing key tasks for a procedure must have a scientific base. It is necessary to demonstrate the construct validity of such models, i.e. that one is measuring the trait which one purports to measure, and this can be inferred by comparison of novice, intermediate and expert performances on the key tasks, assessed by time taken and other measures such as dexterity and error scores (Martin et al., 1997;Joice et al., 1998;Eubanks et al., 1999;Darzi et al., 2001). The metrics from the model must display significant statistical differences in performance between the three subject groups in order to be able to confirm a measurable learning process when used for training. Without demonstration of construct validity, the concept of a learning curve cannot exist as there needs to be a 'skills gap' between novice and expert performance, enabling the former to strive to achieve the skills of the latter.

The validation of the simulated models not only involves the physical model used, but also validation of the assessment parameters used to differentiate between the different groups of practitioners. Assessment of technical skill must be objective, reliable and easy to perform, as further described in section 1.5 for the exemplar of laparoscopic surgery. Nonetheless, the level of technical skills assessment depends upon the task being assessed – it may be appropriate to test simple hand-eye co-ordination tasks solely with dexterity analysis as the procedure is scripted, and thus qualitative differences between subjects are likely to be minimal. However, when fine dissection skills are to be assessed, for example when performing a pelvic dissection to avoid nerve injury, motion tracking alone is not appropriate.

2.3.4 Transfer of Skills to the Real Environment

It is essential to confirm that procedural clinical skills improve when training on a simulated version of a particular task (Seymour et al., 2002;Grantcharov et al., 2004). Achievement of expert performance on the simulated task must, to some extent, transfer to improved skill in the clinical domain. Demonstration of transfer of skills from the laboratory environment to real scenarios has been demonstrated for laparoscopic cholecystectomy and gastrointestinal endoscopy, though one of the main obstacles have been which tools to use for the assessment of real procedures (Scott et al., 2000a;Seymour et al., 2002;Sedlack and Kolars, 2004;Grantcharov et al., 2004).

The same measures developed for assessment in the laboratory environment could be used, and can be further validated by comparison with performance in the operating room, i.e. concurrent validity. Outcome parameters can be defined clinically by error analysis, failed procedure rates or record of complication rates (Van Rij et al., 1995;Tang et al., 2004b;Tekkis et al., 2005). It is crucial to use an assessment tool which captures the desired components of the operative procedure, though to date there is no comprehensive assessment tool for neither laparoscopic nor open surgery.

Traditional outcomes are used as markers of an individual surgeon's technical performance. However, this approach is too simplistic and fails to take account of the numerous factors which can affect patient outcomes (Vincent *et al.*, 2004). Patient characteristics can decrease or increase the risk of complications, especially during major surgical procedures. This can be accounted for through appropriate case selection, ensuring that the sickest or most complex patients are operated upon by the most experienced surgeons. However, it is not only the surgeon who needs to be experienced to ensure an optimal outcome – the rest of the operating team can also have a significant impact on the outcome of the procedure (Undre *et al.*, 2006). This is none more so than for minimally invasive procedures whereby the surgeon must rely upon the camera positioning skills of another individual. For cardiothoracic surgery, an experienced anaesthetist is crucial for those cases where the patient has minimal physiological reserves. Recent work has also shown that post-operative and ward care has a considerable impact upon patient outcomes (Jones *et al.*, 1999).

Furthermore, test cases for evaluation of transfer of skill from the laboratory to the clinical environment require standardisation. This form of case stratification to grade the difficulty of the real procedure can help to ensure that the assessment is of technical skill rather than patient or disease variability (Hanna et al., 1998b;Ragunath et al., 2003). This was not the case for any of the studies regarding transfer of skill from simulated to real clinical environments (Scott et al., 2000a;Seymour et al., 2002;Sedlack and Kolars, 2004;Grantcharov et al., 2004).

2.4 The Impact of a Competency-Based Training Curriculum

Current training programmes in technical skills are primarily run on an ad hoc basis, with the majority organised as two to five day courses (Darzi, 1996). These are intensive, and teach a number of skills to trainees of varying capabilities. Furthermore, a number of the course participants will return to their hospital without further exposure to many of the skills taught on these courses. The development of a structured training programme using a curriculum approach can provide graded teaching sessions. These sessions are targeted toward a specific group of doctors and with pre-set expert-based benchmark criteria to be achieved at every stage of the curriculum, can ensure that everyone in the group benefits from the programme. This increases the motivation of both trainees and faculty involved, providing a more personal approach to technical skills acquisition. The curriculum provides formative feedback to trainees, and can provide a clear marker of those who are falling behind and require further attention from their trainer.

A systematic training and assessment programme is built upon the notion of task-based learning within the skills laboratory, which when transferred to the operative environment also enables deconstruction of a complex procedure into its constituent parts. This provides a structured approach to learning the operative procedure, managing errors and identification of tasks for deliberate practice. The framework also maintains continuity of objective assessment in the real environment, using similar modes as in the skills laboratory. Competency at the real procedure will still entail a learning curve, though the hypothesis to be tested is that this is should be shorter than if traditional modes of training are employed. A comparison of the two methods of training in terms of total time taken, cost and complication rates may provide some evidence to confirm this assumption. Through use of the assessment tools in the real environment, participants in such a program can then be granted certification for independent practice in a technically-based procedure (Dent, 1991;Society of American Gastrointestinal Surgeons, 1991;European Association of Endoscopic Surgeons, 1994).

A more philosophical question though is how we define attainment of competence. The Concise Oxford English Dictionary definition is of one 'having the necessary skill or knowledge to do something successfully' (Soanes and Stevenson, 2004). In medicine, success would be in accordance with the principles of safe practice, to achieve a satisfactory result for each individual patient. Studies of technical skill have generally defined one to be competent in a particular procedure when they have performed greater than 100 of cases of the defined procedure, for example, laparoscopic cholecystectomy. However, this is a poor assumption based upon experience rather than expertise. It is believed that the learning curve has plateaued, though to quote Vince Lombardi (a well known American football coach), 'it is not practice that makes perfect, but *perfect* practice that makes perfect'.

Current assessment of competence fails to be performed in an objective, valid and rigorous manner (Darzi *et al.*, 1999). This concept is being challenged, and achieving a certificate for independent practice is just one point on a spectrum of ability. The ultimate validation of any objective measure of technical skill shall be to correlate this with a clinically relevant outcome measure (Grober et al., 2004b). Once these measures are chosen as the benchmark, the definition of competence can direct the delivery of high quality outcomes. Attainment of this benchmark level of proficiency is a process of summative assessment, though assessment strategies can be further tailored to diagnose and rectify strengths and weaknesses in the process of formative assessment.

This training and assessment continuum can form the educational basis for validation and revalidation of technical excellence. This can augment the development of life-long learning, currently performed through a process of gaining continuing medical education (CME) credits (Ribble *et al.*, 1981). The publication of national databases may be used to provide information regarding outcomes, and provide remedial action as necessary.

Those individuals, or surgical departments, that fail to make the grade should repeat constituent parts of the curriculum, but there is a possibility that some shall continue to lack the skills to become, or remain, competent. The medical profession must deal with this issue head-on by ensuring the scientific base developed from a competency-based training curriculum is openly available to make judgements of this nature. This is only possible be configuring training programmes in a systematic manner as described in this chapter, and can lead to the delivery of high quality care to all patients.

2.5 Application to Non-Medically Qualified Personnel

The framework detailed in this chapter is directed toward the training of technical skills, and may be applied to any technical skill within medicine. Similarly, the framework may also be applied to non-medically trained personnel such as nurses, physiotherapists and paramedics. Intravenous cannulation, application of plaster and endo-tracheal intubation are technical skills which also necessitate a high level of skill to ensure a favourable outcome (Wood, 1994). These skills are commonly learnt on short courses, though again without a scientific base for training and assessment.

The task-based approach to training, commensurate with the educational principles of learning such as concretisation of the task and practice on simulated models in a skills laboratory are applicable here too. A similar process of assessment, both formative and summative in nature can reassure both the patient and practitioner. Utilisation of the same model for training medical and allied health personnel can also encourage the concept of inter-professional learning (Street, 2004). It is not necessary to uphold the hegemonic nature of training in medical tasks, and this framework enables trainee doctors to be supported by experienced nurses when performing skills such as cannulation at the early part of their training (Sweet and Norman, 1995). Furthermore, within the current climate of new professional roles in surgery, a framework of training which is identical for doctors and allied health professionals alike, with the same levels of competence to achieve, can reassure the public and health care community of the continued maintenance of standards of excellence (Duthie et al., 1998;Kneebone and Darzi, 2005).

2.6 Conclusions

Currently, there is a lack of definition toward the methods required for training the health care professionals of the future. Curricula exist, though are based upon experiential notions and are not standardised between regions or centres (Bailey et al., 1991;Bannister et al., 2003;Mertz and Gautam, 2004). This leads to a variation in the quality of training, and more importantly a variation in the definitions of competence for independent practice.

The model charted out in this chapter is simplistic, feasible and generic to any branch of medicine which involves the acquisition of technical skill (figure 2a). The model also provides an opportunity to develop valid, objective and reliable methods to assess technical skill both in laboratory-based and real environments. It is imperative that such an approach is pursued, especially in the context of growing public and political pressures for competency-based practice (Smith, 1998). It will then be possible not only to define levels of skill, but also to audit training programs and provide academic institutions with an objective argument to obtain further resources to fund these programs. Finally, it is also possible that this model of training can reduce the number of unnecessary complications caused by having to learn technical skills on real patients.

Chapter 3

DEVELOPMENT OF AN EVIDENCE-BASED VIRTUAL REALITY TRAINING PROGRAM

Aim

Define a training curriculum on a virtual reality simulator for laparoscopic skills acquisition, which is graded in a stepwise manner, and utilizes valid parameters for assessment of proficiency.

3.1 Introduction

The implementation of a competency-based surgical skills curriculum necessitates the development of tools to enable structured training, with in-built objective measures of assessment (Aggarwal et al., 2004a;Aggarwal et al., 2004b;Reznick and MacRae, 2006). Simulation in the form of virtual reality and synthetic models has been proposed for technical skills training at the early part of the learning curve (Torkington et al., 2000;Satava, 2001b;Gallagher and Cates, 2004). In order to be efficacious, these tools must convey a sense of realism, and a degree of standardisation to enable graded acquisition of technical skills. In accordance to the STATS model (chapter 2), progression along the curriculum is charted by passing pre-defined expert benchmark criteria, which lead onto more technically demanding tasks.

In the laparoscopic era training on inanimate video trainers, and more recently on virtual reality simulators, has been shown to improve skills performance in the operating room (Scott et al., 2000a;Seymour et al., 2002;Sedlack and Kolars, 2004;Grantcharov et al., 2004). The skills required for laparoscopic surgery are markedly different to those employed in open surgery (Figert *et al.*, 2001). It is now generally accepted that these skills must initially be learnt in training laboratories prior to entering the operating theatre. Nevertheless, structured training programs utilising these tools do not exist and have not been validated in terms of which tasks should be performed, at which level, for how long, how often, and to which set of benchmark criteria. Current studies demonstrating the beneficial effect of training on the MIST-VR laparoscopic simulator (Mentice, Gothenburg, Sweden) have employed a variety of training methods, none of which have been evidence-based (Chaudhry et al., 1999;Torkington et al., 2001;Jordan et al., 2001;McNatt and Smith, 2001;Gallagher and Satava, 2002;Hamilton et al., 2002;Kothari et al., 2002;Seymour et al., 2002;Ali et al., 2002;Pearson et al., 2002;Grantcharov et al., 2003;Grantcharov et al., 2004). Previous studies have shown a plateau of the learning curves after the second repetition at the easy level (Ali *et al.*, 2002), and at between the fifth and seventh repetition at the medium level (Gallagher and Satava, 2002;Grantcharov et al., 2003). Though there is a greater paucity of data on the other virtual reality simulator models, similar problems exist. Infact, no studies to date have defined the nature of a structured and proficiency-based training program from the easy level through medium and onto the hard levels.

The aim of this chapter was to develop an evidence-based virtual reality training program for the initial acquisition of technical skill on two laparoscopic virtual reality simulator platforms, leading to a basic level of proficiency prior to entering the operating theatre. Basic and procedural tasks can be simulated in a high-fidelity virtual environment that closely resembles the operative field. Virtual tissues can be manipulated, clipped and cut, and incorporated into a recognisable simulation of Calot's triangle dissection which bleeds and can respond to diathermy (figures 1c, 1d and 1e). At the end of each task, performance can be measured using parameters such as time taken, number of errors made and path length for each hand which are objectively and automatically recorded by the simulator. This makes it possible to chart the performance of a trainee surgeon along a curriculum, and define their attainment of proficiency.

The structured curriculum can enable trainees to be confident in their skills prior to assisting in, and performing the initial laparoscopic procedures, safe in the knowledge that they have achieved pre-set expert criteria. The ultimate aim is to reduce their learning curve on real patients, leading to acquisition of proficiency at an earlier stage than training on patients alone. Airline pilots become proficient at flying an aeroplane before even leaving the ground, acquiring skills on a high-fidelity flight simulator (Roscoe, 1971;Roscoe, 1972). The analogous situation should now be possible for the early part of the learning curve in laparoscopic surgery. This may lead to a reduction in the number of unnecessary complications occurring due to a failure of technical skills (Moore and Bennett, 1995), and the time and expense spent acquiring basic laparoscopic skills in the operating room (Bridges and Diamond, 1999;Babineau et al., 2004).

The MIST-VR training curriculum was the first to be developed (section 3.2), and subsequently data was collected for the development of the LapSim training curriculum (section 3.3).

3.2 MIST-VR training curriculum

The MIST VR system runs on a desktop PC (400MHz Pentium II, 256Mb RAM) with tasks viewed on a 17-inch CRT monitor positioned at operator eye level. The video subsystem employed (Matrox Mystique, 8-MB SDRAM) delivered a frame rate of approximately 15 frames per second, permitting near-real-time translation of instrument movements to the video screen. The virtual laparoscopic interface (or VLI) input device (Immersion Corporation, San Jose, CA) consists of two laparoscopic instruments at a comfortable surgical height relative to the operator, mounted in a frame by position-sensing gimbals that provided six degrees of freedom, as well as a foot pedal to activate simulated electro-surgery instruments. With this system, a 3D 'box' on the computer screen represents an accurately scaled operating space. Targets appear within the operating space according to the specific skill task selected and can be grasped and manipulated with virtual instruments (figure 1c). Each of the different tasks is recorded exactly as performed and can be accurately and reliably assessed.

The first six tasks (core skills 1) have been described in section 1.4.6. In a similar manner, the second six tasks (core skills 2) are described here. All tasks begin with bilateral movements to touch a virtual sphere with the tips of the virtual instruments. For task 1 the trainee is required to stretch a virtual tube and hold it still at the desired length. In task 2, the virtual tube must be clipped in its centre. Task 3 consists of stretching the virtual tube, followed by clipping it in the centre, whilst holding it still at the desired length. Task 4 requires the trainee to stretch a virtual tube to the pre-set length. In task 5, the virtual tube is held still, and it must be diathermied along its length, remaining with the 'safe zone'. Task 6 combines the actions of tasks 4 and 5 with the aim of stretching the virtual tube to the desired length, and then diathermying it along its length, once again attempting to remain within the safe zone.

MIST-VR has a total 12 basic tasks at three levels of difficulty – easy, medium and hard (figure 3.1) (Taffinder et al., 1998;Gallagher et al., 2001;Gallagher and Satava, 2002). Each set of six tasks are complementary, and the sixth task from each group is the most complex, integrating aspects of the previous five. The degree of challenge from easy to hard depends upon the degree of accuracy required by the operator, i.e. the margin of error is smaller at the hard level.

The aim of this study was firstly to define whether there is a relationship between training on the two most complex tasks (i.e. task 6 from core skills 1 and 2) as opposed to training on all twelve MIST-VR tasks, enabling definition of 'key training tasks'. The further aim was to proficiency levels, leading to a step-wise curriculum and proficiency-based for basic laparoscopic skills training.

3.2.1 Methods

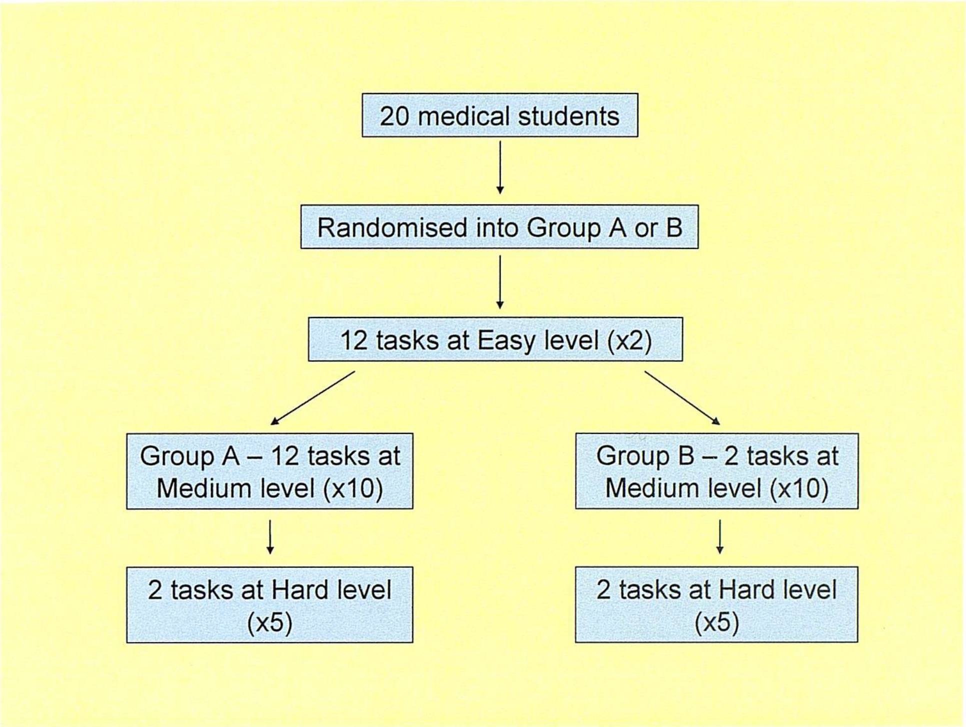
Twenty medical students with no previous laparoscopic experience were recruited to undergo a training program on the MIST-VR simulator. Training sessions were delivered over a maximum period of 14 days, with two sessions per day, each at least one hour apart. A study coordinator was present during all training sessions to provide technical assistance.

To familiarise themselves with the simulator, all participants initially performed two sessions of all 12 tasks at the easy level. This number of repetitions was based upon the previous finding by Ali et al. that the learning curve at the easy level plateaued at the second repetition (Ali *et al.*, 2002). Subjects were then randomised into two equal groups using the sealed envelope technique. Both groups performed ten sessions at the medium level, though Group A continued to train on all 12 tasks whilst Group B only trained only on the two most complex tasks, i.e. tasks 6 from core skills 1 and 2. Finally, all participants completed the two most complex tasks at the hard level for a further five sessions.

Assessment of progress throughout the training programme was measured by the simulator in terms of time taken, economy of movement and number of errors for each hand. In order to maintain uniformity of assessment between the sessions, all comparisons were made on the two most complex tasks only. This enabled comparison between groups A and B, and definition of the number of sessions required prior to flattening of the learning curve for each parameter.

In order to define whether the novices had achieved expert levels of psychomotor skill, their performance at the hard level was compared to that of benchmark scores derived from the performance of 10 expert laparoscopic surgeons on the two complex tasks. The expert criterion was defined as having performed over 100 laparoscopic cholecystectomies in total.

Figure 3a – Study Design



‘12 tasks’ denotes all 12 tasks from core skills 1 and 2;
‘2 tasks’ denotes the two most complex tasks from each of core skills 1 and 2.

3.2.2 Statistical Analysis

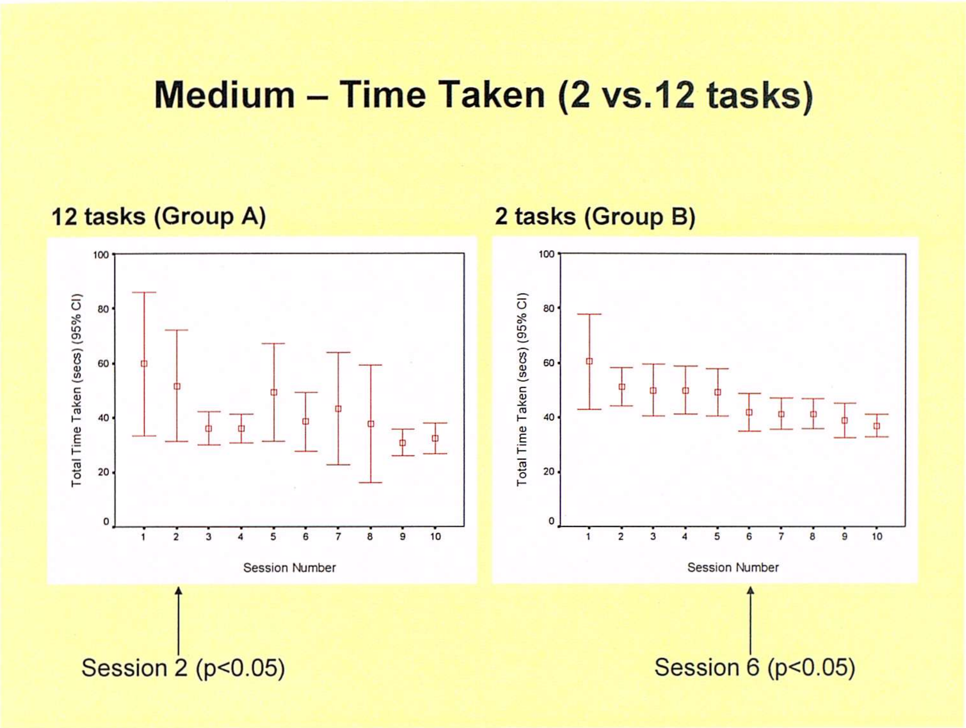
Data was analysed with the Statistical Package for the Social Sciences version 11.5 (SPSS, Chicago, Illinois, USA) using non-parametric tests. Data on learning curves was analysed by the Friedman (non-parametric repeated measures ANOVA) test. Multiple comparisons were then made to identify when plateau of skills had occurred. Comparison of performance between groups A and B was undertaken using the Mann-Whitney U Test. A level of $p<0.05$ was considered statistically significant.

3.2.3 Results

All 20 participants completed the study curriculum, and there were no demographic differences between groups A (12 tasks) and B (2 tasks). At the easy level, baseline performance scores were similar for groups A and B in terms of time taken ($p=0.588$), economy of movement ($p=0.813$) and error scores ($p=0.496$).

Assessment of performance during training at the medium level for groups A and B on task six for core skills 1 revealed a statistically significant flattening of the learning curve at sessions two and six respectively for all parameters of time taken, error scores and economy of movement (figures 3b, 3c and 3d). Similar scores were achieved for task 6 on core skills 2. At the medium level, there were no significant differences when comparing skill levels achieved by group A at the second session with those achieved by group B at the sixth session.

Figure 3b – Time Taken at the Medium Level



Medium – Time Taken (2 vs.12 tasks)

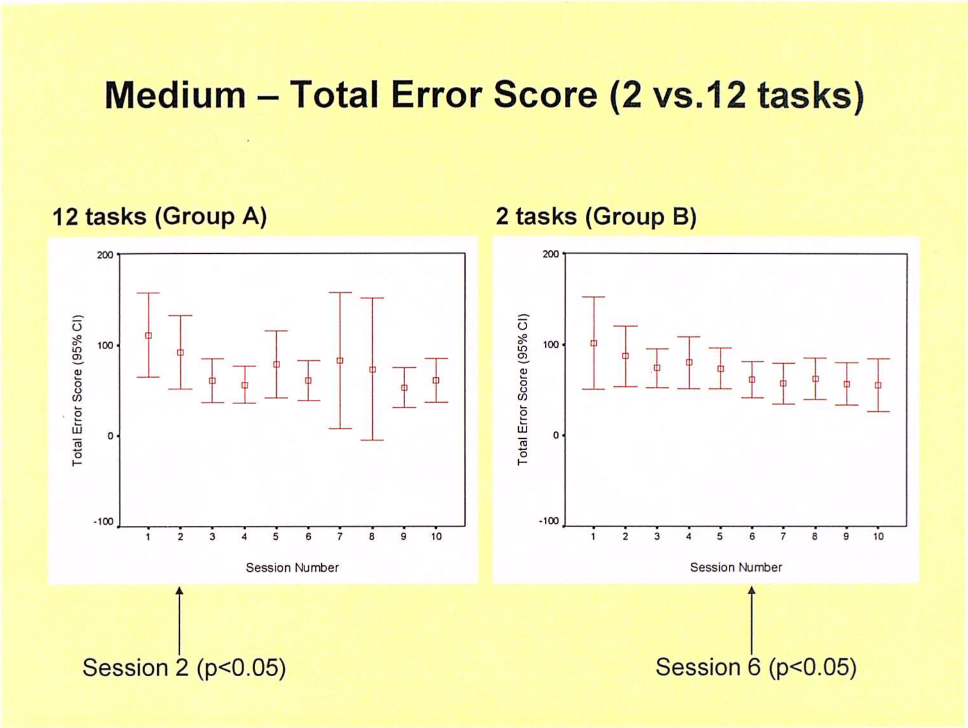
12 tasks (Group A)

2 tasks (Group B)

Session 2 (p<0.05)

Session 6 (p<0.05)

Figure 3c – Total Error Scores at the Medium Level



Medium – Total Error Score (2 vs.12 tasks)

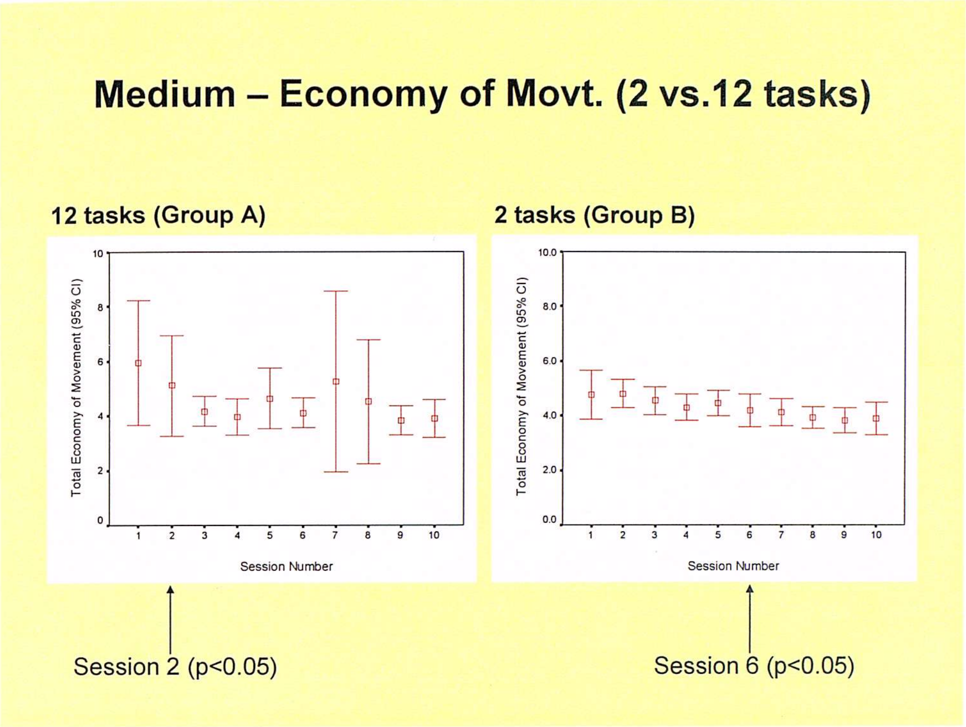
12 tasks (Group A)

2 tasks (Group B)

Session 2 ($p<0.05$)

Session 6 ($p<0.05$)

Figure 3d – Economy of Movement at the Medium Level



Medium – Economy of Movt. (2 vs.12 tasks)

12 tasks (Group A)

2 tasks (Group B)

Session 2 (p<0.05)

Session 6 (p<0.05)

At the hard level, there was no evidence of significant improvement in skill for either group A or B (p>0.05), and again there were no differences in skill levels between the groups for economy of movement (p=0.152) or error scores (p=0.914). However, group A were significantly faster than group B (p=0.006) at completing the tasks at the hard level. The median scores of 10 experienced laparoscopic surgeons were taken as benchmark levels and comparison of the participants' scores with these benchmark levels revealed that all students were able to achieve the set criteria at the hard level.

3.2.4 Application of results to a MIST-VR training curriculum

This study has provided information for the establishment of evidence-based virtual reality training curriculum for the acquisition of psychomotor skill. Previous studies have demonstrated plateau of skill acquisition at the easy and medium levels (Gallagher and Satava, 2002;Ali et al., 2002;Grantcharov et al., 2003). However, it was the aim of this study to define a stepwise curriculum moving from easy to medium and onto hard levels of difficulty, and to provide objective evidence regarding the most efficient use of the simulator.

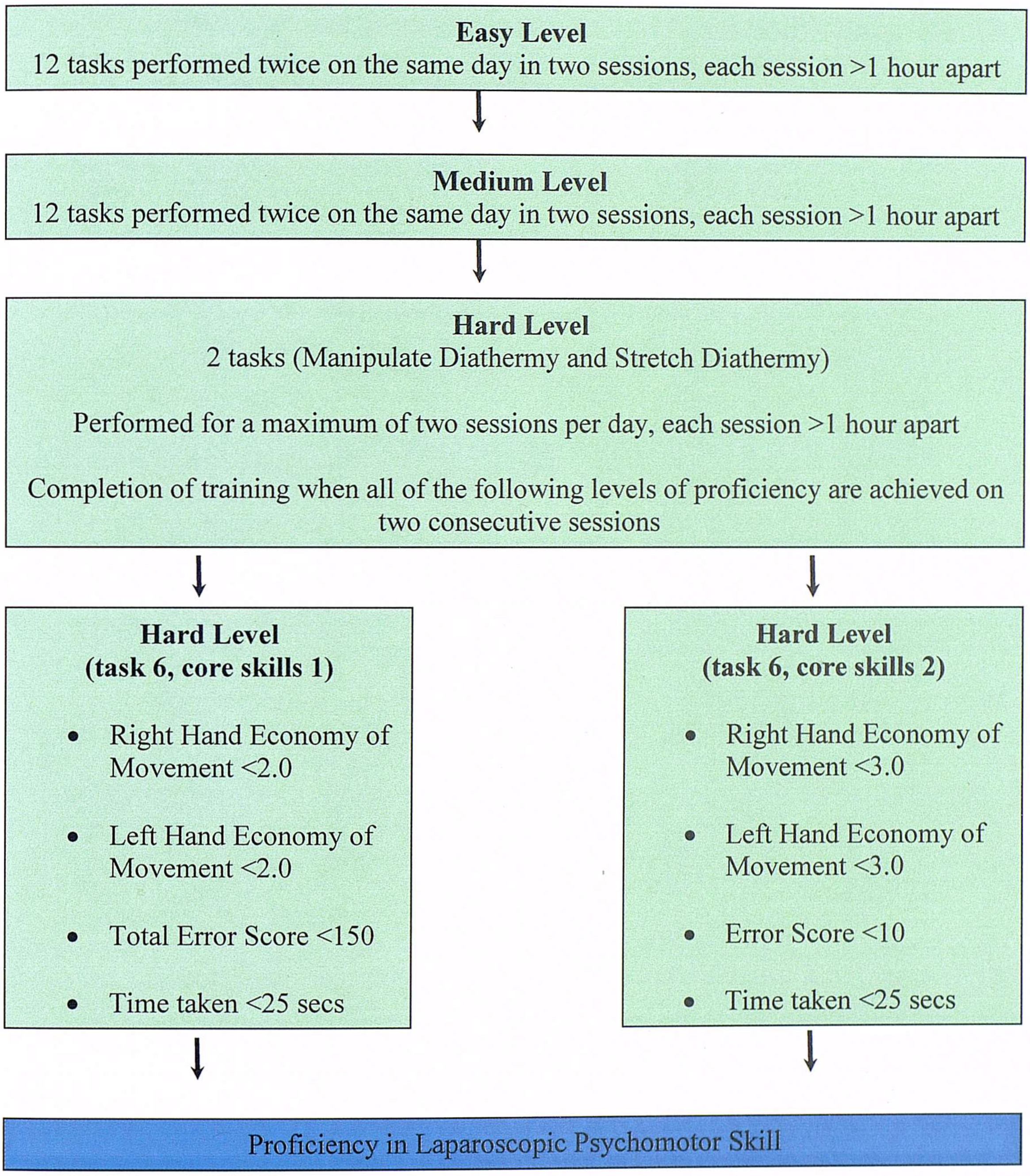
This study has shown that it is possible to acquire the same level of skill whether one performs only the two most complex tasks, or all twelve tasks on the simulator, i.e. the tasks are indeed complementary. Subjects in the 12 tasks group took approximately 30 minutes to complete each session, whereas subjects in the 2 tasks group took about 5 minutes per session. It is evident that a greater number of sessions are required when only two tasks are practiced. Regardless of the training curriculum adopted at the medium level, all subjects in both groups managed to attain expert levels of skill when performing at the hard level.

Integrating the results of this study into a training curriculum can be achieved by alignment with a framework that describes three stages of motor skills learning proposed by Fitts and Posner (Fitts AM and Posner MI, 1967). During the first stage (cognition), the trainee gains an understanding of the task through instructor explanation and demonstration. This relates to the quality of the human-machine interaction, resulting in a short phase for a realistic model. In the second stage (association), the user practices the task to eliminate error, with the instructor providing feedback to identify errors and suggest corrective actions. The final phase (automation) occurs when the learner performs the tasks in a relatively automated fashion with little or no cognitive input. Extrapolating the concept to this study, it can be seen that training at the easy level enables the trainee to familiarise themselves with the controls of the simulator and the nature of the tasks to be performed. This is in tandem with another study commenting on the high quality user-computer interface of the MIST-VR simulator (Gor *et al.*, 2003). The medium level provides an opportunity to refine performance at the tasks, leading onto the hard level by which time the skills have become automated, or second nature. This is evident not only by the fact that there was no further improvement in performance at the hard level, but also that all subjects had achieved expert benchmark levels of skill too.

It is thus possible to define this three-stage curriculum as one that enables trainee surgeons to familiarise, train and be assessed on a virtual reality simulator (figure

3e). Assessment of skill can be performed at the hard level by comparison with the benchmark values of an expert group of surgeons. The results of this study confirm that all participants achieved levels of skill similar to those of a group of experienced laparoscopic surgeons. By increasing the numbers of experienced surgeons in the database, it will ultimately be possible to test the predictive validity of these criteria for a trainee surgeon's pathway to achievement of competence (Gallagher *et al.*, 2003b). It is also important to note that the benchmark levels of skill must be achieved at two consecutive sessions, to confirm acquisition of skill rather than attainment of a good score by chance. Finally, a maximum of two sessions are allowed per day, performed at least one hour apart. This can ensure a distributed rather than massed approach to skills training (Mackay *et al.*, 2002).

Figure 3e – A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill.



Easy Level

12 tasks performed twice on the same day in two sessions, each session >1 hour apart

Medium Level

12 tasks performed twice on the same day in two sessions, each session >1 hour apart

Hard Level

2 tasks (Manipulate Diathermy and Stretch Diathermy)

Performed for a maximum of two sessions per day, each session >1 hour apart

Completion of training when all of the following levels of proficiency are achieved on two consecutive sessions

Hard Level
(task 6, core skills 1)

- Right Hand Economy of Movement <2.0
- Left Hand Economy of Movement <2.0
- Total Error Score <150
- Time taken <25 secs

Hard Level
(task 6, core skills 2)

- Right Hand Economy of Movement <3.0
- Left Hand Economy of Movement <3.0
- Error Score <10
- Time taken <25 secs

Proficiency in Laparoscopic Psychomotor Skill

3.3 LapSim training curriculum

The LapSim virtual reality simulator used in this study creates a virtual laparoscopic system that consists of a computer operating system (Windows XP, Microsoft, Redmond, Washington, USA), a video monitor and a virtual laparoscopic interface (or VLI, same as for the MIST-VR platform) containing two pistol-grip instruments and a diathermy pedal. Neither the MIST-VR or LapSim platform which were used had the ability of force (or haptic) feedback.

There are seven basic tasks, of which clipping, grasping and cutting have already been described (section 1.4.3). The remaining tasks are: camera navigation, co-ordination, instrument navigation, and lift and grasp. In camera navigation the trainee has to focus on virtual spheres placed around the operative scene, before they disappear. Co-ordination requires the use of one hand to control the camera to find the virtual spheres, and the other to touch them with a virtual instrument. Instrument navigation is a task whereby successive virtual spheres appear which must be touched without damage to the tissue background. Finally, lift and grasp is a task whereby a virtual box has to be lifted and a virtual needle grasped from underneath the box; the needle is then to be placed into a virtual 'cup'. All tasks can be performed at varying degrees of complexity according to the following settings: object size, target size, object or target timeout, moving camera, and number of objects per task. There is also the possibility of using a 30 degree, as opposed to a zero degree, endoscope.

At time of the study, there was one procedure-specific task which comprised dissection of Calot's triangle. The trainee has a wide choice of laparoscopic instruments (toothed and non-toothed grasper, straight and curved dissector, scissors and suction) all of which can have the added benefit of diathermy if so desired. The task requires the subject to safely and accurately dissect the cystic duct and artery, clip each, and divide them. The amount of inflammatory tissue placed over the cystic structures can be varied, together with bleeding rate. Adjacent structures such as the common hepatic artery and bile duct are also present, and may be damaged.

3.3.1 Methods

The study recruited 40 participants of whom 19 were experienced laparoscopic surgeons (performed >100 laparoscopic cholecystectomies) and 21 were inexperienced (<10 laparoscopic cholecystectomies). Twenty subjects were chosen at random using the closed envelope technique, though stratified to comprise 10 experienced and 10 novice laparoscopic surgeons. This group of surgeons performed only the basic skills modules (i.e. tasks 1 to 7). The remaining 20 subjects (i.e. 9 experienced and 11 novice surgeons) performed only the dissection module. The nature of the study was explained to all subjects prior to enrolment, and informed consent was obtained in all cases.

3.3.1.1 Tasks and Skill Levels

The LapSim® virtual reality laparoscopic simulator (Surgical Science, Gothenburg, Sweden) has seven basic tasks, at three different levels of difficulty, i.e. easy, medium and hard. The tasks aim to teach skills for instrument navigation, handling tissues and clip application. The default settings at each level are defined arbitrarily through a default setting on the software, though a previous study has confirmed construct validity of the assessment parameters at the easy level (Duffy *et al.*, 2005). To incorporate training sessions at the medium and hard level, the settings used must be construct valid, i.e. measure the traits one purports measure, which can be inferred by showing differences in performance between experts and novices. To ensure acquisition of skills, repeated practice must lead to an improvement in the performance of trainees, and plateau toward that of the experienced surgeons. The pre-set expert criteria to achieve can be defined by analysing the scores of a pool of experienced laparoscopic surgeons on the simulator.

Since there were no data at time of study design to confirm the construct validity of the medium and hard levels, this was the aim of this study. However, the validity of each level would necessitate a new group of subjects, and in the interests of efficiency it was decided to only concentrate upon the validity of the hard level. As the easy level had been previously determined to be valid, if the settings of the medium level lay between those of the easy and hard level, it could be suitably assumed that the settings of the medium level were also construct valid.

The default settings at the hard level of the simulator were difficult, though unrealistic. For example, many modules incorporated an awkward camera to target angle, which if in a realistic situation would be altered by repositioning of the camera or port sites. The settings for the tasks performed at the medium and hard level were thus redefined prior to commencement of the study (example shown in table 3a).

This was performed independently by two of the study investigators, leading to a consensus definition for the settings of the medium and hard settings. Subjects (i.e. 10 expert and 10 novice laparoscopic surgeons) performed the seven tasks on the basic skills module, solely at the hard level, over 10 sessions. No more than three sessions were completed per day, each at least one hour apart to ensure tiredness was not a factor during the sessions.

The settings for the procedural (Calot's triangle dissection) module were once again defined by two study investigators prior to commencement of the study. Verification of this module as a valid tool for training and assessment was undertaken in a similar manner to that for the basic skills modules. Nine experienced surgeons performed two repetitions of the module, on the same day. Similarly, 11 novices performed 10 repetitions of the module, with no more than three sessions per day.

Table 3a – Example settings for the basic skills modules at easy, medium and hard for 'lift and grasp'

SETTINGS	**HARD**	**MEDIUM**	**EASY**
Scope Angle	0	0	0
Moving Camera	Y	N	N
No. of Left Balls	3	3	3
No. of Right Balls	3	3	3
Object Timeout	12 secs	12 secs	None
Object Size	20mm	20mm	20mm
Target Size	28mm	28mm	40mm

3.3.1.2 Performance Evaluation

Comparison of median performance at the first two sessions between the two groups of surgeons can assess whether each simulated task is construct valid, and substantiates the use of the defined settings of the simulator to assess laparoscopic technical skill. Statistical analysis of the learning curves for novice laparoscopic surgeons clarifies whether repeated practice improves performance toward that of the experienced group.

The definition of benchmark criteria to be achieved prior to progression onto the next stage of the curriculum was by calculation of the median score for each parameter (which was defined to be construct valid) over the two sessions performed by all 10 experienced surgeons. This provided 20 expert data points for each task, and a consensus of operative performance for novice surgeons to attempt to achieve. Assessment of the experienced surgeons' first two repetitions also enables familiarisation and reduces the effect of the early learning curve on the benchmark scores.

3.3.2 Statistical Analysis

The choice of 10 subjects in each group was based upon power calculations from previous studies on virtual reality simulation. The data was analysed with the Statistical Package for the Social Sciences version 11.5 (SPSS, Chicago, Illinois, USA) using non-parametric tests. Data on learning curves was analysed by the Friedman (non-parametric repeated measures ANOVA) test. Multiple comparisons were then made to identify when plateau of skills had occurred.

Comparison of performance between expert and novice groups was undertaken using the Mann-Whitney U Test. A level of $p<0.05$ was considered statistically significant.

3.3.3 Results

All seven basic tasks demonstrated construct validity ($p<0.05$) at the hard level for the parameters of time taken and path length. Error scores which include instrument misses, 'tissue damage' and 'stretch damage' did not validate for any of the tasks, except for accuracy of clip placement on the 'clip applying' task.

Analysis of the learning curves for novice subjects on all tasks revealed significant improvements in validated parameters at a median of seven repetitions, though individually the 'lift and grasp' and 'clip applying' tasks had the longest learning curves. The experienced laparoscopic surgeons demonstrated learning curves on all tasks at a median of the second repetition, with the longest curves once again on 'lift and grasp' and 'clip applying' tasks. A graphical representation of the learning curves of experienced and novice surgeons for time taken, instrument path length and error scores during repetitions of the 'clip applying' task are shown (figure 3f).

On the procedural module (Calot's triangle dissection), time taken, blood loss, dissected volume, path length and angular path length were significantly different ($p<0.05$) between experienced and novice surgeons over the first two repetitions, confirming the construct validity of this module as an assessment tool. Learning curves for the 11 novice subjects on the validated parameters plateaued at the fourth repetition (range 2nd – 5th repetitions).

Figure 3f – Learning curves on 'Clip and Cut' task of LapSim virtual reality simulator for (i) time taken, (ii) instrument path length and (iii) error scores. (Legend: ▲ = Experienced group; ■ = Inexperienced group)

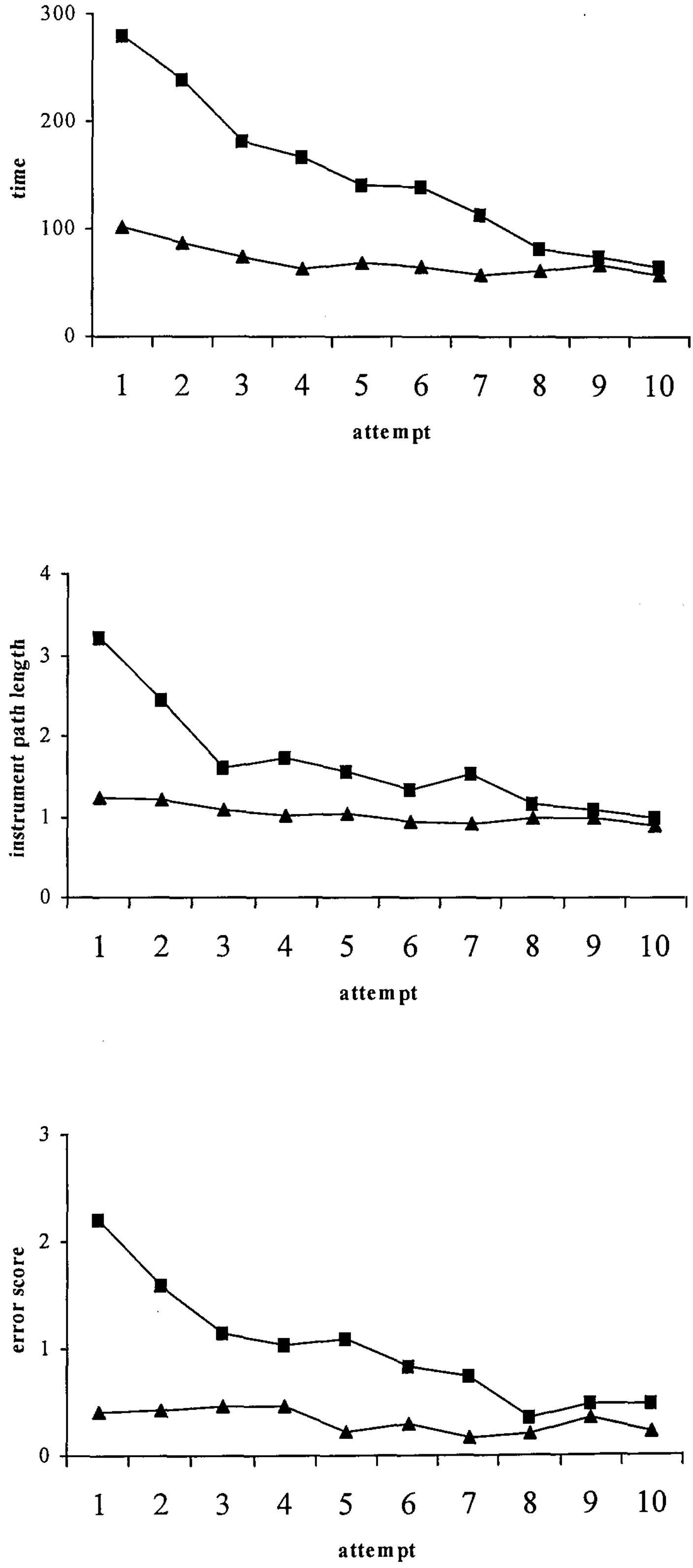
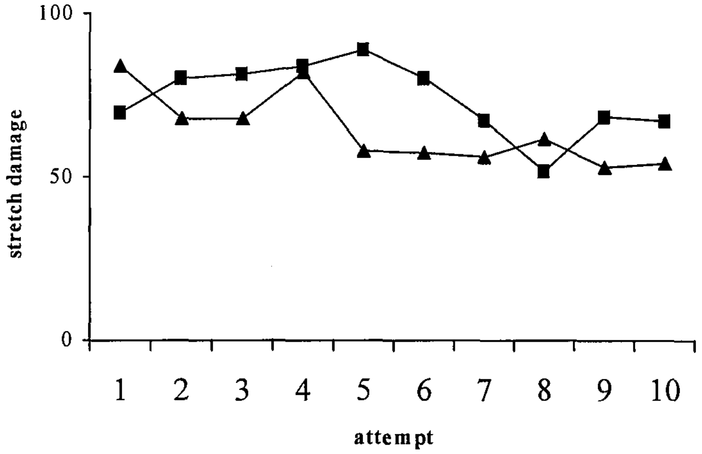
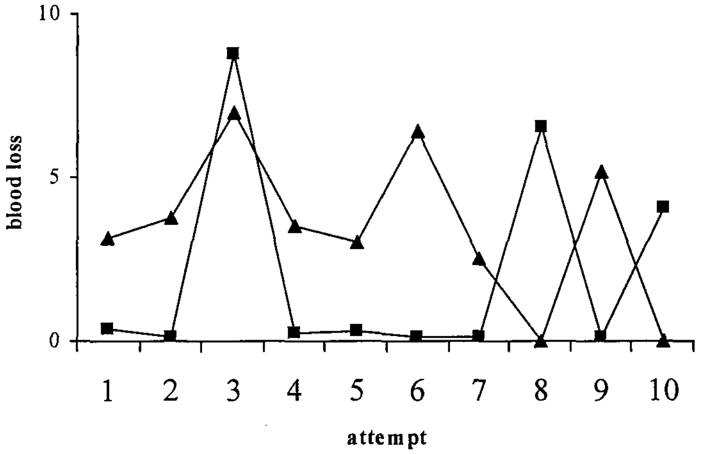


Figure 3g – Learning curves on 'Clip and Cut' task of LapSim virtual reality simulator for (i) stretch damage and (ii) blood loss.
(Legend: ▲ = Experienced group; ■ = Inexperienced group)





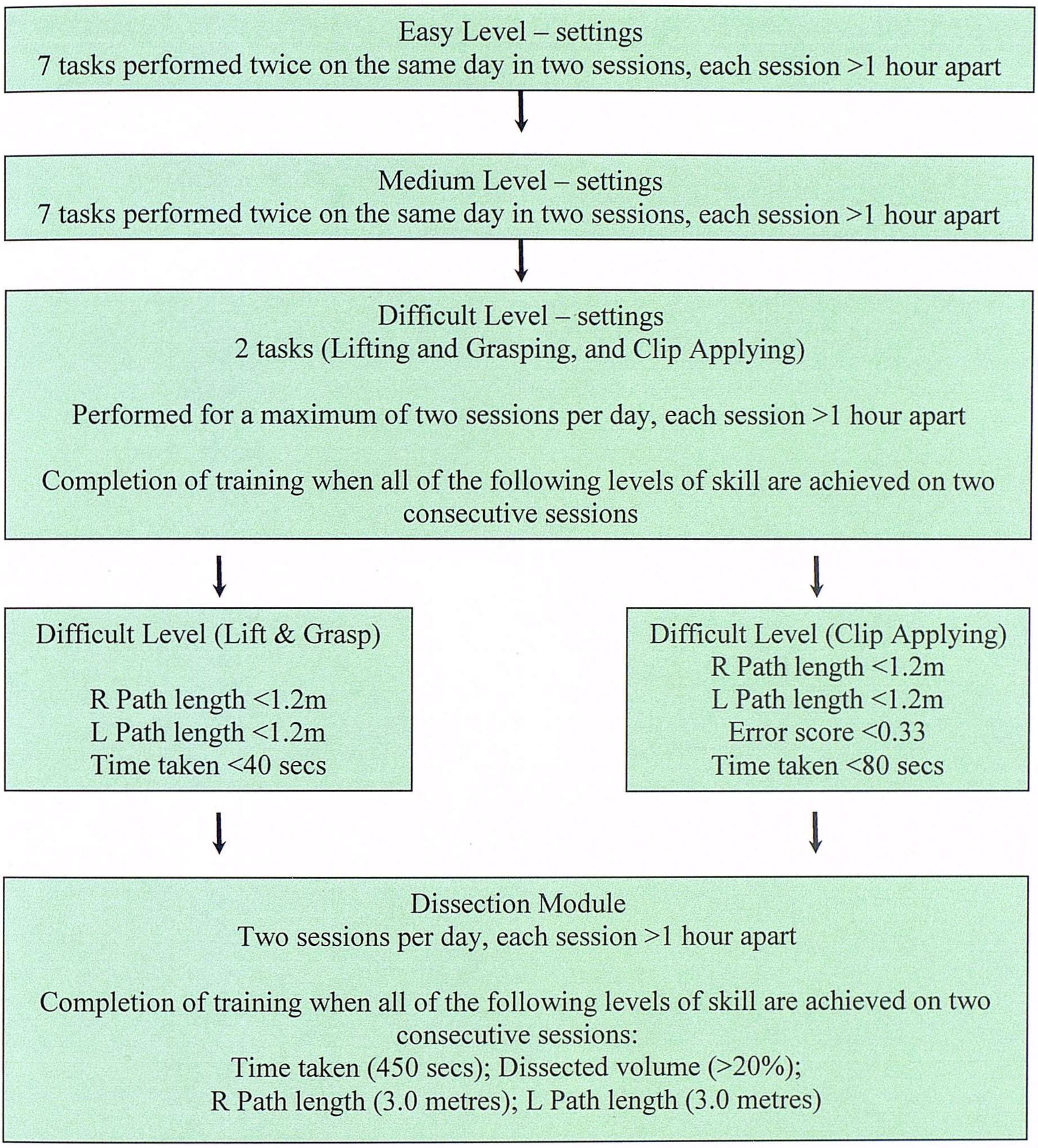
3.3.4 Application of results to a LapSim training curriculum

The results of this study enable the definition of a model of training on this simulator, which can then become one of the components of a competency-based training curriculum for laparoscopic surgery (figure 3h). Data from the previous MIST-VR curriculum study has shown that commencement of training at the easy level on all basic tasks for two sessions, followed by equivalent training at the medium level for a further two sessions is the most appropriate method for initiation of laparoscopic skills training (section 3.2).

At the hard level, the sessions involve training at the two most challenging tasks, ‘lift and grasp’ and ‘clip applying’. These were the tasks with the longest learning curve for both novice and experienced subjects, echoing the inference of McDougall et al. on the LapMentor simulator (McDougall *et al.*, 2006). It is thus appropriate that these are the two test tasks which are chosen for demonstration of proficiency. Benchmark criteria were thus defined for ‘lift and grasp’ and ‘clip applying’ tasks on the validated parameters of time taken, path length and error score (figure 3f). Learning curves are also shown for two of the parameters which did not achieve construct validity, i.e. stretch damage and blood loss (figure 3g).

Progression onto the procedural stage of the training program i.e. dissection module, (following achievement of benchmark criteria on the basic skills module), entails a further two sessions per day, repeated until the pre-set expert benchmark levels have again been attained. As for the MIST-VR training curriculum, it is important to note that the benchmark levels of skill must be achieved at two consecutive sessions, to confirm acquisition of skill rather than attainment of a good score by chance. Finally, a maximum of two sessions are allowed per day, performed at least one hour apart. This can ensure a distributed rather than massed approach to skills training (Mackay *et al.*, 2002).

Figure 3h – An evidence-based virtual reality training program for novice laparoscopic surgeons



3.4 Conclusions

A competency-based training curriculum for laparoscopic surgery must utilise valid tools which enable the trainee to practice on a series of standardised technical tasks (Aggarwal *et al.*, 2004a). Pre-defined benchmark criteria can ensure skills acquisition has been successful (Korndorffer, Jr. et al., 2005;Katz et al., 2005). A procedure-based technical skills curriculum must begin by teaching the basic skills necessary for laparoscopic surgery such as hand-eye coordination, the fulcrum effect and depth perception (Rosser *et al.*, 1997). Subsequent procedural skills training enables integration of knowledge and judgement into the technical skills already learnt (Fried *et al.*, 2004b).

The curricula charted out in this chapter have been developed through a process of scientific validation, and incorporates a competency-based approach. Individual tasks at individual levels of difficulty have been proven to be construct valid, primarily upon the basis of quantitative parameters, i.e. time taken and economy of movement. Learning curve analysis proves that novice surgeons improve their performance with repeated practice on the simulator, toward that of experienced surgeons. The technical skills taught by training on the simulator are thus relevant for laparoscopic surgery, and the supposition is that this should lead to a shortening of the learning curve on real patients in the operating room (Roscoe, 1971;Roscoe, 1972). It is also important to note that only the parameters which achieved construct validity on the basis of scores during the first two sessions displayed a clear learning curve, i.e. there is no point repeating a task if there is no clear 'skills gap' between expert and novice surgeon at commencement of training (figure 3f and 3g).

This leads onto the question as to why the majority of the error-based assessment parameters such as tissue damage did not validate. It may be due to a lack of realism of the dynamic characteristics of the virtual tissues, or more likely because of inherent difficulties in objectively defining a surgical error (Cuschieri, 2003). There is in fact no standard definition for medical error, though a working definition is given by Reason: 'an error is defined as a generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome' (Reason, 1994). It is relatively simple for a computer to measure the quantitative parameters such as amount of instrument movement, in an accurate manner, though the occurrence of an activity which has failed to achieve its intended outcome is difficult even for the human brain to quantify. The only applicable manner in which to accurately quantify errors is to ensure surgeons perform to a pre-defined protocol, with clear definitions of intended and adverse outcomes (Joice *et al.*, 1998). However, surgery is not a mechanical process, and thus it is difficult to wholly justify this course of action.

It can be argued that the curriculum charted for these two virtual reality simulators is at present applicable only to the simulator, and not to real procedures. The key point though is that it is achievement of pre-set criteria, rather than the number of sessions or repetitions, that determines completion of the training period. It is thus irrelevant how long one takes to complete the program, and some trainees shall of course attain proficiency faster than others

(Korndorffer, Jr. et al., 2005;Katz et al., 2005). However, all trainees who have achieved the pre-set criteria should have the psychomotor skills necessary to assist in, and learn to perform, real laparoscopic procedures. The curricula charted out provide a prescriptive approach to training novice laparoscopic surgeons, to an evidence-based level of proficiency in basic psychomotor skills.

This model of skills training enables the development of proficiency in basic laparoscopic skills. Some centres have sought to define the requirements of basic and advanced laparoscopic training programs, with a move away from two-day courses toward an extended curriculum over a period of weeks to months (Powers et al., 2002;Coleman and Muller, 2002;Lin et al., 2003). However, the structure of these programs has been developed in a largely arbitrary manner rather than through scientific evaluation.

The curricula such as those presented within this chapter can become the nuts and bolts of a validated and competency-based laparoscopic training program incorporating the use of video trainers, VR simulators, porcine models and ultimately training in the operating theatre. Prior to adding further modes of surgical skills training though, it is necessary to document the role of virtual reality simulation within a surgical curriculum. The virtual reality curricula presented must be shown to lead to significant improvements in laparoscopic technical skill in the clinical setting, i.e. show transfer of training. Grantcharov et al. and Seymour at al. have proven this, though not with a pre-defined virtual reality training curriculum, rather with an un-validated and non-evidence-based training program (Seymour et al., 2002;Grantcharov et al., 2004). This is the continued aim of this research thesis, though a fundamental problem to be tackled prior to evaluation of the 'transfer of training' of these curricula is to develop a feasible, objective, valid and reliable mode of assessment within the operating theatre.

Chapter 4

AN EVALUATION OF THE FEASIBILITY, VALIDITY AND RELIABILITY OF LAPAROSCOPIC SKILLS ASSESSMENT IN THE OPERATING ROOM

Aim

Investigate the feasibility, validity and reliability of laparoscopic skills assessment in the operating theatre which utilizes motion tracking and video-based analysis to assess technical performance.

4.1 Introduction

Recent publications of the rates of medical errors and adverse events within healthcare, and particularly during surgery, have drawn the spotlight toward the methods of credentialing surgeons to perform procedures independently (Reason, 1994;Smith, 1998;Cuschieri, 2003;Elwyn and Corrigan, 2005). Training boards and certifying bodies are coming under increasing pressure to ensure individuals demonstrate the necessary skills to perform operations safely (Ribble et al., 1981;Dent, 1991;Society of American Gastrointestinal Surgeons, 1991;European Association of Endoscopic Surgeons, 1994;Jackson B, 1999). This is not only important for patient safety, but also underpins the development of a proficiency-based training curriculum.

It is somewhat surprising then that there are no tools in widespread use which are feasible, valid and reliable for assessment of technical surgical skill (table 4a) (Moorthy *et al.*, 2003). Current training outcomes are assessed by live evaluations of the trainee by the master surgeon, a process that is known to be biased and subjective (Cuschieri *et al.*, 2001). More objective data is available from morbidity and mortality data, though this is rarely a sole function of operative skill and thus does not truly reflect an individual's surgical competence (Vincent *et al.*, 2004). The majority of trainees also maintain a log of the procedures performed, but these are indicative merely of procedural performance rather than a measure of technical ability (Reznick, 1993;Reznick and MacRae, 2006).

Table 4a – Qualities of the ideal surgical assessment tool

Feasibility		is a measure of whether something is capable of being done or carried out
Validity	Face validity	is the extent to which the examination resembles real life situations
	Content validity	is the extent to which the domain that is being measured is measured by the assessment tool—for example, while trying to assess technical skills we may actually be testing knowledge
	Construct validity	is the extent to which a test measures the trait that it purports to measure. One inference of construct validity is the extent to which a test discriminates between various levels of expertise
	Concurrent validity	is the extent to which the results of the assessment tool correlate with the gold standard for that domain
	Predictive validity	is the ability of the examination to predict future performance
Reliability	Test-retest	is a measure of a test to generate similar results when applied at two different points
	Inter-rater	is a measure of the extent of agreement between two or more observers when rating the performance of an individual

Though a number of new tools have been developed to assess surgical technical performance, their use remains within the confines of surgical skills laboratories (Moorthy et al., 2003;Aggarwal et al., 2004b). These include virtual reality simulators and psychomotor training devices which are designed primarily to assess performance during critical parts of a procedure, rather than a complete operation (Aggarwal *et al.*, 2004b). The realism (or face validity) of such simulations is not perfect and the situations lack context, leading to a failure of operators to treat the models like real patients (Kneebone *et al.*, 2003).

The ideal device for objective assessment of real surgical procedures would be one that can automatically, instantly and objectively provide feasible, valid and reliable data regarding performance within the operating room (Darzi *et al.*, 1999).

It is with this approach that our Department has developed the Robotic Video Motion Analysis System (or ROVIMAS) motion tracking software which enables surgical dexterity to be quantified and thus reported instantly by a computer program (Dosis *et al.*, 2005). Though automatic, objective and instant, the motion tracking data does not provide any information regarding the quality of the procedure performed. ROVIMAS does however incorporate the ability to synchronously record video of the operative procedure, which can then be evaluated according to a valid and reliable rating scale (Martin et al., 1997;Joice et al., 1998;Eubanks et al., 1999). This can enable a definition of dexterity not only for whole procedures, but also for critical steps of a particular procedure.

A preliminary publication has confirmed the feasibility of using the device within the operating room to assess laparoscopic skills (Dosis *et al.*, 2005). The primary aim of this study was to determine the validity and reliability of a new concept for technical skills assessment in the operating room – a combination of motion analysis and video assessment. Both approaches have been individually validated in the literature, though the introduction of a hybrid between the two modes of assessment has not been previously attempted. If successful, then the application of different operative training schedules or curricula can be objectively, reliably and validly assessed.

4.2 Methods

4.2.1 Subjects

Nineteen surgeons were recruited to the study, and subdivided into six novice (<10 laparoscopic cholecystectomies, LCs) and thirteen experienced (>100 LCs) practitioners. The aim was for each surgeon to perform a minimum of two procedures (to determine inter-test reliability), with consecutive cases recorded over a period of six months.

4.2.2 Operative Procedure

Laparoscopic cholecystectomy was chosen as the operative procedure of choice as it is a common operation, performed in a fairly standardised manner and amenable to both motion tracking and video-based analysis. Not only LC is an index procedure for commencement of training and ongoing assessment of laparoscopic skills, but there is great interest in the current use of simulation-based devices to improve surgical technical performance in the operating theatre (Dent, 1991;Society of American Gastrointestinal Surgeons, 1991).

4.2.3 Patients

Ethical approval was obtained from the local research ethics committee to record video data of each operative procedure. Patients were recruited from two surgical departments and all were consented prior to entry into the study. The aims of the study were explained, and patients were free to withdraw from the study at any time. Furthermore, it was made clear that the only intervention was to record the intra-operative procedure. No distinguishing features were to be recorded and all tapes were to be coded alphanumerically rather than by name or date of procedure.

In order to reduce the effect of disease and patient variability, all patients recruited to the study were deemed to have a diagnosis of biliary colic. To enable objectification of this approach, inclusion and exclusion criteria were developed with subsequent classification according to patient and disease state, and post-hoc review of the video tape according to operative state (table 4b) (Hanna *et al.*, 1998b).

Table 4b – Inclusion and exclusion criteria for patient entry to the study

		Inclusion criteria	**Exclusion criteria**
Patient Characteristics	*Age*	>18yrs and <65 yrs	<18 yrs and >65 yrs
	Obesity	BMI < 30kg/m^2	BMI > 30 kg/m^2
	Anaesthetic risk	ASA = 1 or 2	ASA > 2
	Hospital admission with gallbladder pathology	No	Yes
Disease Characteristics	*Diagnosis*	Biliary colic	Acute cholecystitis
	Complications of gallstones	None	Any
	ERCP	No	Yes
	Blood tests (at any time pre-operatively)	WCC < 11 CRP < 5 LFTs – normal range	WCC > 11 CRP > 5 LFTs - abnormal
	Ultrasound findings (at any time pre-operatively)	Gallstones / sludge	Thickened gallbladder wall Peri-cholecystic fluid Ultrasonographic Murphy +ve Common bile duct stone Common bile duct dilatation
Intra-operative characteristics (modified from Hanna et al.)	*Degree of difficulty*	Cystic duct seen on retraction of gallbladder	Contracted, inflamed, or densely adherent gallbladder Gallbladder neck adherent to bile duct
		Unobstructed view of Calot's triangle, or Fat over Calot's triangle	Fat-laden falciform Hypertrophied liver: quadrate lobe partially obstructing view and/or right hepatic lobe making retraction difficult
		No obvious ductal or vascular anomaly	Difficult, obscure, abnormal anatomy
		None or filmy/loose areolar adhesions to gallbladder	Dense omental adhesions to gallbladder Duodenal adhesions to gallbladder
			Stone impacted in neck or Hartmann's pouch

4.2.4 ROVIMAS assessment device

All procedures were recorded with the ROVIMAS software which synchronously recorded video and motion analysis data from the operative procedure. The purpose of the study was explained to all surgeons prior to the patient consent process. Once scrubbed, surgeons wore one pair of sterile gloves over their hands. Sensors were then placed onto the dorsum of each hand, followed by the donning of surgical gown and a further pair of sterile gloves. This avoided the need to sterilise the electromagnetic sensors, and friction between the gloves maintained the sensors in the correct position. The procedure was demonstrated by one of the study investigators, who remained present throughout the surgical procedure. Once the patient had been anaesthetised, the electromagnetic emitting device was placed onto their sternum, fixed firmly by Micropore tape (3M Corporation, St. Paul, MN, USA).

The video feed from the laparoscopic stack was recorded onto the ROVIMAS software through a digital video link (I-link, IEEE-1394) to a laptop computer. Recording commenced upon entry of the endoscopic camera into the peritoneal cavity and was complete upon removal of the camera from the abdomen for the final time. The open parts of the procedure were not recorded, the aim being to solely assess the laparoscopic skills of the subjects. Complete, unedited videos of each procedure were recorded with the software into Microsoft Windows .avi format (Microsoft Corporation, Redmond, WA, USA). All data files were coded by an alphanumeric code to ensure the identity of the operating surgeon and patient were blinded to the reviewers.

4.2.4.1 Motion analysis device

The parameters used for this study were those which have been previously validated on bench-top assessments of laparoscopic skill on a porcine model, i.e. time taken, path length and number of movements for each hand (Smith *et al.*, 2002).

4.2.4.2 Video-based assessment

The objective structured assessment of technical skill (OSATS) proposes a generic evaluation of surgical performance through use of a global rating scale (Martin *et al.*, 1997). The scale was initially validated through live-marking of bench-top tasks and is said to 'boast high reliability and show evidence of validity' (Dath *et al.*, 2004). The aim was to determine the validity, inter-rater and inter-test reliability of this scale for assessment of laparoscopic technical skills in the operating room. Rating on the scale was performed by two experienced laparoscopic surgeons who were blinded as to the identities of the operating surgeons.

4.3 Data collection and analysis

Three dimensional co-ordinate data from the Isotrak II motion tracking device (Polhemus Inc, Colchester, VT, USA) was translated into useful parameters of time taken, path length and number of movements of each hand by the ROVIMAS software (see Dosis et al. for a detailed description (Dosis *et al.*, 2005)). With the aid of the synchronisation feature of ROVIMAS, data was derived for the entire procedure, and also for pre-defined parts of the procedure, classified into tasks as shown in table 4c. It must though be noted that the values for the whole procedure are not a sum of all the parts identified, i.e. insertion of accessory ports, division of adhesions, removal of gallbladder, etc.

Power analysis was based upon results from a previous study on motion analysis of porcine LCs, and revealed a sample size of 20 cases per group(Smith *et al.*, 2002). Statistical analysis employed non-parametric tests of significance. Construct validity was determined by comparison of performance between novice and experienced surgical groups for dexterity parameters and scores from video-rating scales with the Mann-Whitney test. The Cronbach's alpha test statistic was used to ascertain the inter-rater reliability of the video-based scoring system. Inter-test reliability was assessed by comparison of the first and second consecutive procedure performed by each surgeon, once again with the Cronbach's alpha test statistic.

Table 4c – Definitions of the tasks of a laparoscopic cholecystectomy

	Start	**Finish**
Whole procedure	First moment of insertion of endoscopic camera	Final removal of endoscopic camera
Dissection of Calot's triangle	The first moment gallbladder is grasped at Calot's triangle	Entry of clip applicator to the operative field of view
Clip and cut cystic duct	First entry of clip applicator to the operative field of view prior to clipping the cystic duct	Cystic duct is clipped and divided
Clip and cut cystic artery	First entry of clip applicator to the operative field of view prior to clipping the cystic artery	Cystic artery is clipped and divided
Dissection of gallbladder from liver bed	Following division of duct and artery, the first moment that the peritoneum between gallbladder and liver bed is grasped	Gallbladder is freed from liver

In order to investigate the existence of a relationship between dexterity analysis and the video-based rating scale for assessment of surgical performance, correlations between the two methods were calculated with the non-parametric Spearman's rank correlation coefficient. For all tests, $p < 0.05$ was considered statistically significant.

4.4 Results

A total of 53 procedures were performed by the 19 surgeons recruited to the study. Six cases were excluded in an independent manner by both reviewers upon the basis of intra-operative characteristics (table 4b). Of the remaining 47 cases, 14 were performed by six novice surgeons and 33 by the 13 experienced surgeons. The median number of cases carried out by each surgeon was 2 (range 1–5 cases).

4.4.1 Motion tracking data (table 4d)

A comparison between LCs performed by novice and experienced surgeons revealed significant differences in time taken for the whole procedure (median 2175 vs. 1979 seconds, p=0.036), though not for total path length or number of movements. This result was replicated for 'clip and cut duct' (55 vs. 33 seconds, p=0.013) and 'clip and cut artery' (37 vs. 21 seconds, p=0.004).

Only the task concerning dissection of Calot's triangle produced significant differences between the performance of novice and experienced surgeons for all three motion tracking parameters (figures 4a-c): time taken (854 vs. 393 seconds, p=0.002), total path length (138 vs. 73 metres, p=0.026) and total number of movements (640 vs. 367, p=0.005).

Table 4d Results from motion tracking parameters, divided into individual tasks. Values are given as medians according to each group, with the inter-quartile range in parentheses. p values are based upon inter-group comparisons from the Mann-Whitney test. Values in coloured type are those which were statistically significant.

	Time taken (seconds)			**Total path length (metres)**			**Total number of movements**		
	Novice (n=14)	*Experienced (n=33)*	*p value*	*Novice (n=14)*	*Experienced (n=33)*	*p value*	*Novice (n=14)*	*Experienced (n=33)*	*p value*
Whole procedure	2175 (1954-3127)	1979 (1137-2582)	0.036	440 (391-565)	423 (274-667)	0.625	1708 (1599-2072)	1771 (1015-2303)	0.389
Dissection of Calot's triangle	854 (768–1056)	393 (243-691)	0.002	138 (107-196)	73 (39-167)	0.048	640 (528-866)	367 (197-583)	0.007
Clip and cut cystic duct	55 (40-105)	33 (18-57)	0.013	8 (3-15)	5 (2-7)	0.063	23 (14-42)	21 (10-35)	0.553
Clip and cut cystic artery	37 (27-104)	21 (13-30)	0.002	4 (3-12)	3 (2-5)	0.119	18 (8-33)	13 (8-17)	0.204
Dissection of gallbladder from liver bed	401 (233-837)	374 (207-620)	0.471	75 (61-141)	69 (51-132)	0.377	351 (250-761)	325 (227-541)	0.396

Figure 4a – Time taken for experienced and novice surgeons to dissect Calot's triangle. There was a significant difference between experienced and inexperienced groups (p=0.002)

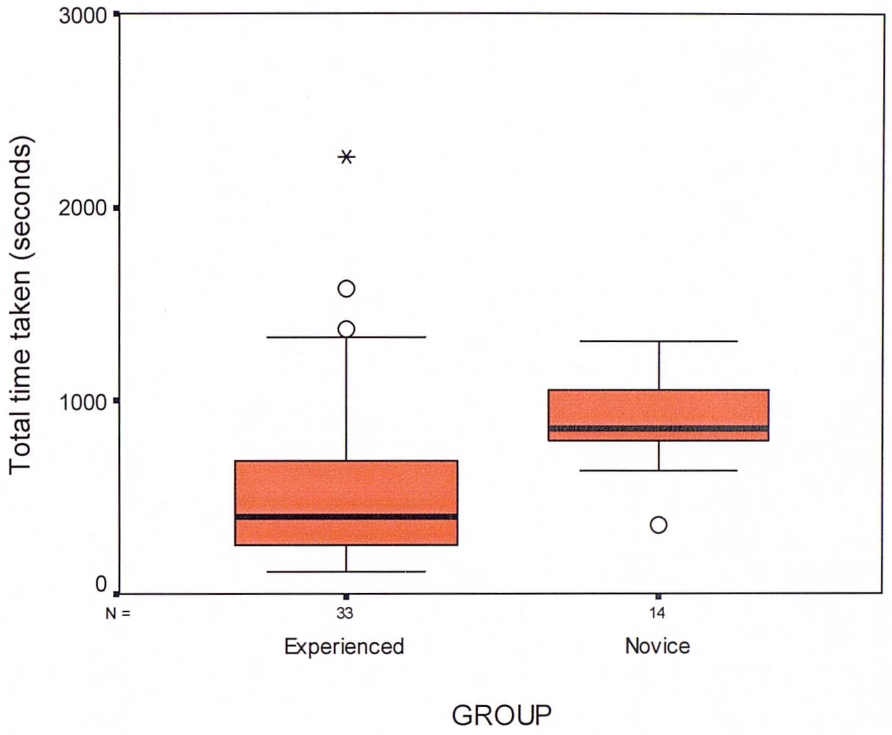


Figure 4b – Total path length for experienced and novice surgeons to dissect Calot's triangle. There was a significant difference between experienced and inexperienced groups (p=0.048)

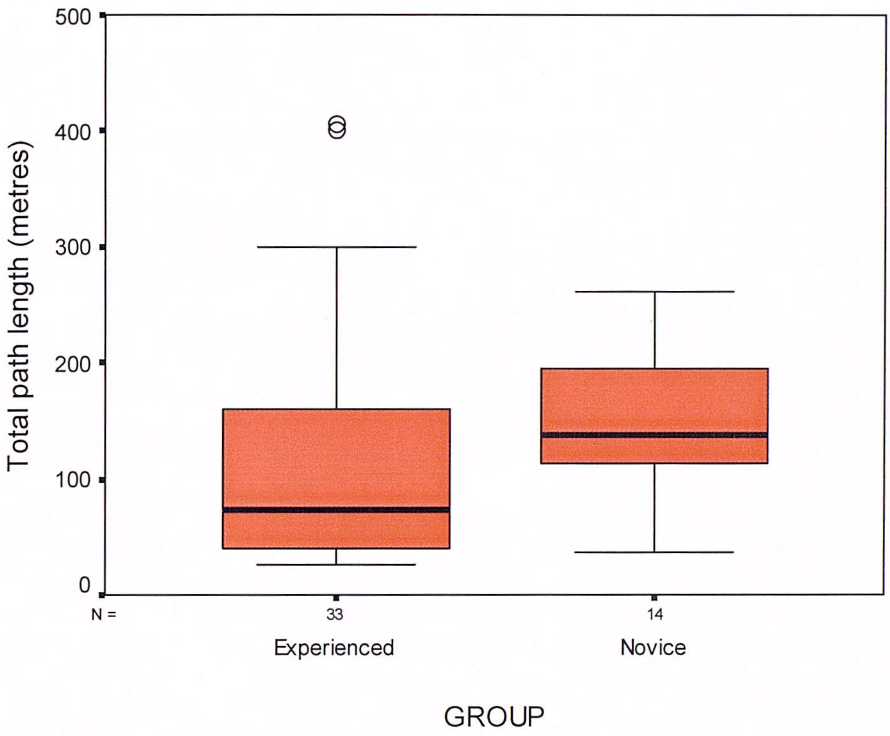
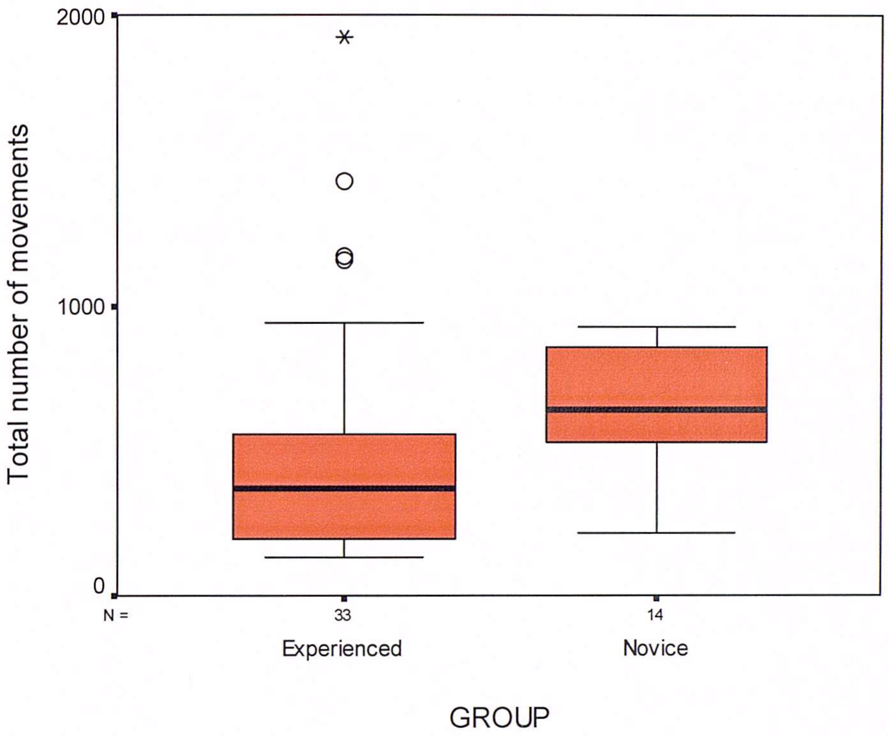


Figure 4c – Total number of movements for experienced and novice surgeons to dissect Calot's triangle. There was a significant difference between experienced and inexperienced groups (p=0.007)

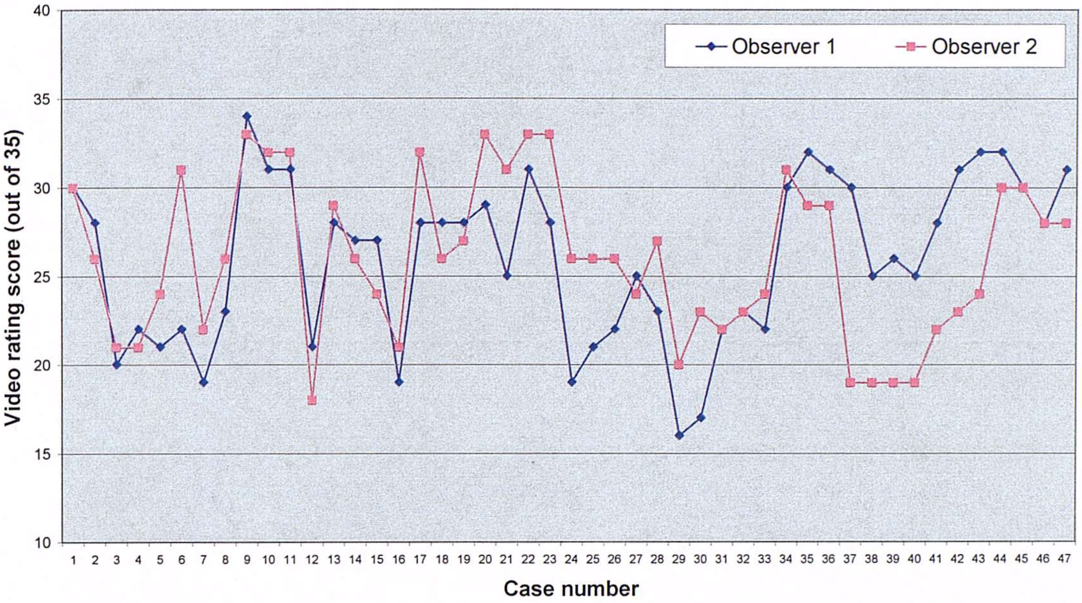


Fifteen of the nineteen surgeons performed two or more procedures each. The inter-test reliability between their first two consecutive cases for time taken to perform the whole procedure was α=0.502. With regard to dissection of Calot's triangle, inter-test reliability was calculated for time taken (α=0.623), total path length (α=0.229) and total number of movements (α=0.522).

4.3.2 Video-based data

The generic OSATS global rating scale demonstrated a significant difference in scores between the novice and experienced surgeons (median 24 vs. 27, p=0.031), with an inter-rater reliability coefficient of α=0.72. The inter-test reliabilities of the fifteen surgeons who performed two or more procedures for the OSATS was α=0.72.

Figure 4d – Graphical representation of inter-rater reliability of OSATS global rating score between observer 1 and 2 (Cronbach's $\alpha = 0.72$)



4.3.3 Comparison of motion tracking and video-rating scale

The correlations between scores obtained from the OSATS global rating scale and validated motion tracking parameters are shown in table 4e. All r values were statistically significant, and ranged from between 0.4 to 0.7, indicating that there were good correlations between the two modes of assessment.

Table 4e Correlation between OSATS global rating score and motion tracking data, calculated using Spearman rank correlation test

		Correlation with video score (r value)	**Correlation with video score (p value)**
Whole procedure	Time taken (seconds)	-0.625	p<0.001
Calot's triangle	Time taken (seconds)	-0.468	p<0.001
	Total path length (metres)	-0.411	p<0.005
	Total number of movements	-0.414	p<0.004
Clip/cut duct	Time taken (seconds)	-0.603	p<0.001
Clip/cut artery	Time taken (seconds)	-0.625	p<0.001

4.5 Conclusions

Though surgical competence is a multi-modal function, proficiency in technical skills to perform an operative procedure is fundamental to a successful outcome (Spencer, 1978;1991;Hall et al., 2003). Assessment within the operating theatre is not only a mode of credentialing individual surgeons, but also enables collective audit of surgical units and residency training programs (Urbach and Baxter, 2004). Despite the development of a number of tools for assessment of technical skills, none have been incorporated into standard practice (Moorthy *et al.*, 2003). This is either due to their complexity, poor validity or the lack of experienced personnel to administer them. The only way to ensure data is collected for every single operation performed within a hospital is to develop a system which automatically records and analyses the required information, without causing any delay or difficulty to the operating room procedure.

This was the intention with the development of the ROVIMAS motion tracking system. With this study it has been possible to reiterate the feasibility, and confirm the validity and inter-test reliability of this device for assessment of laparoscopic technical skills in the operating theatre. In terms of validity, time taken was the only marker to differentiate whole procedure performance between groups of experienced and novice laparoscopic surgeons. However, the synchronisation feature of the system enabled motion tracking parameters to reveal significant differences in dexterity between the two groups of surgeons during a specific part (or task) of the operative procedure, namely dissection of

Calot's triangle. The novice surgeons on average were twice as slow and half as dextrous as the experienced group. This may be because this is the most difficult part of the operation, and indeed the most likely to lead to a catastrophic error.

Standardisation of the procedures was performed on the basis of the patient history, pre-operative investigations and intra-operative findings (Hanna *et al.*, 1998b). A closer inspection of this process reveals that standardisation is primarily based upon the degree of inflammation at Calot's triangle. It is thus not surprising that this part of the procedure gleaned significant differences in dexterity between the two groups of surgeons. None of the other parts of the operation demonstrated construct validity for assessment of dexterity parameters. Reasons for this are either due to simplicity of the task e.g. clip and cut, or anatomical variations such as length of the gallbladder attached to the liver bed. It may also be possible that the failure to achieve significance in dexterity parameters for the whole procedure and other parts of the operation between the two groups is due to an underpowered study. Though the intended 20 cases per group were recruited, six were excluded upon the basis of intra-operative characteristics.

Nonetheless, it is of concern to note that there remained similar degrees of variability within the novice and experienced groups in terms of dexterity assessments. It would be expected, and indeed has been shown in the literature, that experienced surgeons display a greater degree of consistency when compared to their junior counterparts (Mackay *et al.*, 2002). The conflicting factor may be that surgeons performed the procedure with their 'usual technique', perhaps explaining the variability within the experienced group. It would be necessary to confirmed this by dexterity analysis of different techniques to perform laparoscopic cholecystectomy, e.g. blunt/sharp versus blunt/teasing dissection methods (Joice *et al.*, 1998).

A significant limitation of dexterity-based assessment using motion analysis is a failure to capture the qualitative and procedural aspects of an operation. Though number of movements can be used as a measure of operative dexterity, performance is more readily measured by number of faulty or inappropriate movements. This was the reason for integration of a video-based analysis of technical skill, in this case using the OSATS global rating scale (Martin *et al.*, 1997). Though other rating scales exist, divided broadly into checklists and error-scoring systems, the OSATS global rating scale has repeatedly been validated for skills assessment. In this study, the generic OSATS scale displayed construct validity.

The aim of a global rating scale is to assess general surgical principles, whereas checklist-based assessments are by definition specific to the operation. Checklists enable detailed evaluations by specifying individual steps and sub-steps of the operative procedure (Cao *et al.*, 1999). This is time-consuming in terms of assessment, and awards the surgeon only if they perform the procedure in the pre-defined sequence of steps. However, surgery is not a mechanical process and thus it is difficult to justify its evaluation in such a rigid manner. The criticism is that such as scale can only ensure whether something was done or not, but not whether it was done well or poorly.

A number of studies have made use of the generic OSATS scale within the operating theatre (Regehr et al., 1998;Scott et al., 2000a;Scott et al., 2000b;Grantcharov et al., 2004), though the architects of the tool themselves are concerned that 'critical aspects of technical skill are not assessed' (Dath *et al.*, 2004). Furthermore, this scale was developed for use in live rather than video-based assessment. The Toronto group have subsequently developed and validated a video-based, procedure-specific objective component rating scale (OCRS) for Nissen fundoplication (Dath *et al.*, 2004). In the same vein, researchers have recently sought to develop procedure-specific global rating scales, though their efficacy is yet to be tested (Larson et al., 2005;Sarker et al., 2005).

Joice et al. developed an error-based approach to surgical skills assessment utilises human reliability analysis (HRA), which is the systematic assessment of human-machine systems and their potential to be affected by human error (Joice *et al.*, 1998). Tang et al. have applied an Observational Clinical Human Reliability Assessment (OCHRA) tool to laparoscopic cholecystectomy and pyloroplasty (Tang et al., 2004a;Tang et al., 2004b;Tang et al., 2005;Tang et al., 2006). Observational videotape data is subjected to a detailed step-by-step analysis of surgical operative errors, which are divided into consequential or inconsequential. This is a highly specialised and labour intensive task, though it has been suggested that OCHRA provides a comprehensive objective assessment of the quality of surgical operative performance by documentation of errors, stage of the operation when they are most frequent, and when they are consequential. In a not dissimilar manner, the use of dexterity analysis has enabled identification of Calot's triangle dissection as part of the operation in which there are significant differences between novice and experienced surgeons. However, OCHRA benefits assessment in a more formative manner, enabling error modes to be studied and corrective actions to be pursued. Part of the further work to expand this project is to define the relative roles of motion analysis, rating scales and the OCHRA tool for surgical skills assessment.

A further question which has rarely been discussed in terms of surgical skills research is the inter-test reliability of an instrument for assessment of technical skill (Bann *et al.*, 2005). The reliability of the assessment was good for motion tracking parameters and excellent for the video-based global rating scale. This adds further weight to the use of these modes of assessment to assess improvements in performance of a trainee surgeon, or indeed the consistency of an experienced surgeon.

It seems that reliable and valid assessment of laparoscopic skills within the operating theatre can be performed with ROVIMAS motion tracking software, though not exclusively. The global rating scales were valid, and demonstrated higher inter-test reliabilities. It is not surprising then that a comparison of motion tracking data with video-based rating scales revealed significant correlations with the global rating scales. An assumption is that the automated motion tracking device can do the work of two experienced observers assessing an operation on a global rating scale. Though not true at present, this is certainly a notion worth working toward.

With regard to dissection of Calot's triangle, significant differences were noted for *both* path length and number of movements. Though both parameters are related, it is of course possible to perform a task with very fine movements, leading to a large number of movements and shorter path length. Current work seeks to define the relationship between these two parameters, i.e. average path length per movement – a low ratio would suggest fine movements. This information may be useful to highlight accuracy, or perhaps uncertainty, during video-based assessment of the surgical procedure.

Observer-based assessment of technical skill is time-consuming, and relies upon the availability of experienced surgeons to rate performance (Dath *et al.*, 2004). The dexterity parameters from the motion tracking device may be useful as a first-pass filter, avoiding the need to view the entire procedure in order to obtain information regarding technical proficiency. In this manner, surgeons could be automatically assessed on motion analysis each time they performed a procedure, and parts of the video rated either by global or error-based scoring systems only if their dexterity parameters fell outside a pre-determined range of values. This could reduce the time taken to assess proficiency, and would lead to the development of an accurate record of operative skill. Furthermore, many surgeons already make a video record of laparoscopic procedures which they perform; the association of motion tracking data could be stored in a similar operative library.

Extending this concept further leads onto the notion of an operating room black box (ORBB) whereby all aspects of performance are recorded onto a single platform (Guerlain *et al.*, 2005). As well as a documentation of the surgeon's technical skills, the record would include information regarding the rest of the operative team, the patient, equipment used, etc. As in the airline black box, all of this information could be recorded in an automated manner and available for review at a later stage.

Chapter 5

PROVING THE EFFECTIVENESS OF VIRTUAL REALITY SIMULATION FOR TRAINING IN LAPAROSCOPIC SURGERY

Aim

Prove the effectiveness of a virtual reality training curriculum in improving laparoscopic skill on real procedures, in terms of transfer of skill to real procedures, evaluation of the learning curves between simulation-trained and control groups, and measurement of the transfer-effectiveness ratio of the simulator-based curriculum to real procedures.

5.1 Introduction

It is now generally accepted for technical skills training to commence in the skills laboratory on simulated tissues (Moorthy et al., 2003;Fried et al., 2004a;Reznick and MacRae, 2006). These range from simple knot-tying boards through to high-fidelity virtual reality simulators for minimally invasive procedures. Recent studies have shown simulation-based training to transfer to improved performance on real cases, in terms of reduced time taken, fewer errors and decreased patient discomfort (Seymour et al., 2002;Sedlack and Kolars, 2004;Blum et al., 2004;Grantcharov et al., 2004). The overall aim is to shorten the length of the learning curve when acquiring skills to perform procedures on patients. This should not only be a more cost-effective method of training, but also lead to enhanced levels of patient safety.

The over-arching benefits of training in an educationally orientated environment which enables review of performance and the ability to make errors without consequences are powerful incentives to employ simulators within the medical curriculum (Aggarwal *et al.*, 2004a). Though there is a great deal of interest in simulation-based training, especially for the high-fidelity virtual reality (VR) simulators, their presence remains within the confines of research departments, with a paucity of data to prescribe the widespread application of such tools (Sutherland *et al.*, 2006).

Although the intention is to shorten learning curve on real procedures, no studies to date have objectively investigated the degree to which virtual reality simulators in medicine satisfy this demand. Within the airline industry, the method of establishing the quality of a new simulator is to assess its *transfer-effectiveness ratio* (Roscoe, 1971;Roscoe, 1972). The difference in number of trials or time taken to achieve performance criterion (in the air) between untrained and simulator-trained pilots is divided by total training time received by the simulator-trained group. It is thus possible to calculate how cost and time-effective the addition of a new simulator would be in a training program (Rantanen and Talleur, 2005;Taylor et al., 2005).

The aim of this chapter was to investigate the ability of an evidence-based virtual reality training curriculum for acquisition of laparoscopic skills to firstly, transfer to improved performance (which is objectively assessed) on real laparoscopic procedures, and secondly to determine the transfer-effectiveness ratio of the simulator-based curriculum (chapter 3.3) (Aggarwal *et al.*, 2006a). With this information, it should be possible to accurately chart the effects of integrating simulation-based training programs into the medical curriculum. The two-fold approach would seek to define initial outlay costs, and then to determine the time taken to recover those costs by projected reductions in time and expense spent learning laparoscopic skills in the operating theatre (Bridges and Diamond, 1999). Only with such data can we warrant the use of simulator-based training as part of standard practice.

5.2 Methods

5.2.1 Subjects

The study recruited 30 participants of whom 10 were experienced laparoscopic surgeons (performed >100 laparoscopic cholecystectomies [LCs]) and 20 were novice surgeons (performed zero but observed >5 LCs). The 20 novice laparoscopic surgeons were randomly allocated to either control or VR training groups using the closed envelope technique.

5.2.2 Training Program for Novice Surgeons

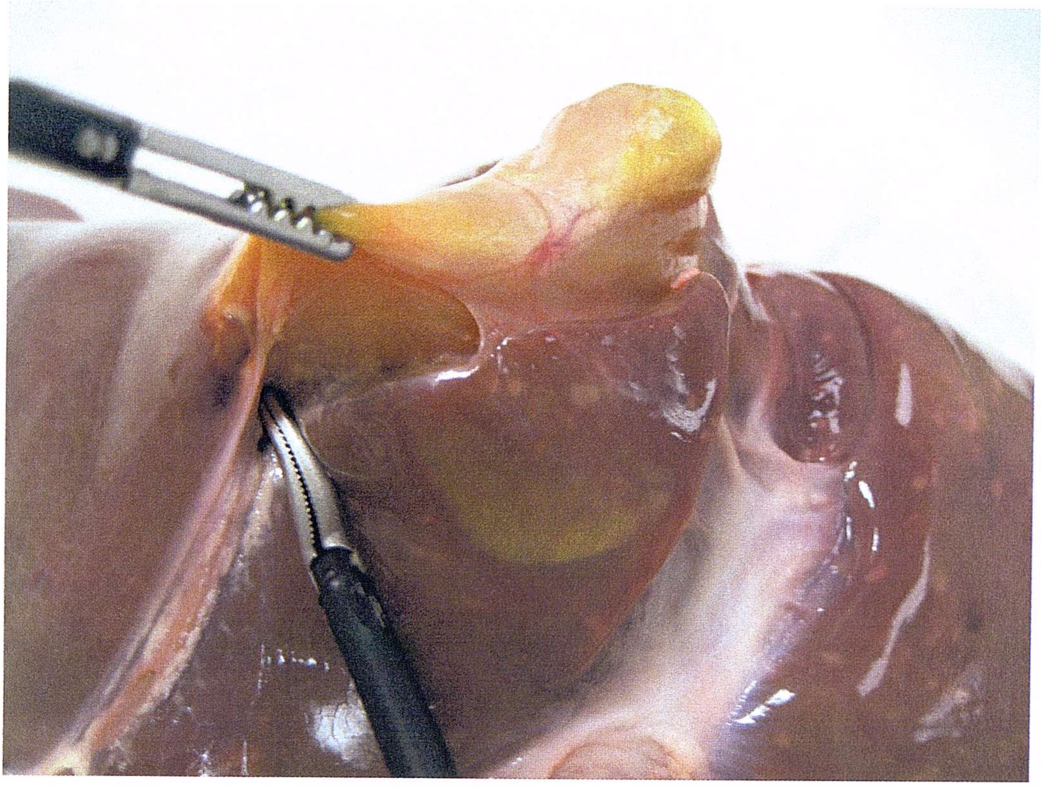
The 20 novice laparoscopic surgeons initially underwent a test of their baseline laparoscopic skills using a previously validated 'glove model' task placed in a box trainer (figure 5a) (Kumar *et al.*, 2003). Performance was measured using dexterity parameters derived from the use of a validated motion tracking device (Smith *et al.*, 2002;Dosis *et al.*, 2005), and number of errors scored in a blinded manner on the end-product by two experienced raters. The intention was to ensure that there were no differences in baseline laparoscopic skill between those randomised to the control or VR trained groups.

Figure 5a – Laparoscopic glove model for basic skills assessment

All novice subjects subsequently attended a one-day laparoscopic skills training seminar during which they were introduced to the concepts of laparoscopic surgery, the different instruments available, and inherent difficulties in using them. A step-by-step approach to laparoscopic cholecystectomy was described with illustrations and video footage of live human operations adapted from the Royal College of Surgeons of England training course (Darzi, 1996). The subjects were then shown videos of two porcine laparoscopic cholecystectomies, one performed by an experienced surgeon and the other by a novice surgeon (figure 5b). The purpose of this was to make the subjects aware of best practice, and perhaps more importantly of errors which may occur during the procedure together with strategies to correct them (Cuschieri, 2003;Seymour et al., 2004).

According to the randomisation process, half of the novice subjects were allocated to training on the VR training curriculum as described below, and the other half allocated to control, i.e. no further training. The control group directly proceeded onto performing five porcine cadaveric LCs over a maximum period of four weeks. The VR trained group underwent the training curriculum until they achieved the prescribed proficiency levels. Following this, they performed three porcine cadaveric LCs each, again over a maximum period of four weeks.

Figure 5b – Snapshot of the porcine laparoscopic cholecystectomy



5.2.2.1 Virtual reality simulator training curriculum

The Lapsim® virtual reality laparoscopic simulator (Surgical Science, Gothenburg, Sweden) has seven basic tasks at three levels of difficulty (easy, medium and hard) to teach skills for instrument navigation, grasping tissues and clip application (figure 1e). A further procedural module comprises one task during which the subject must complete dissection of Calot's triangle. The validity of the tasks and the subsequent development of a training curriculum with proficiency-based criteria to achieve has been previously described (chapter 3.3).

5.2.2.2 Cadaveric porcine laparoscopic procedures

Blocks of cadaveric porcine liver with gallbladder attached were placed into a video-box trainer in an anatomical position, as prescribed on standardised training courses (figure 5b) (Darzi, 1996). A four-port technique was used with ports placed in standardised positions upon the synthetic abdominal wall. Video recording of the procedure commenced from entry of the endoscope into the box trainer through to complete detachment of the gallbladder from the liver bed. The table and display screen height was adjusted to ensure optimal ergonomic conditions for all subjects (Emam *et al.*, 2002). Subjects were provided with standard laparoscopic instruments: straight grasping forceps (toothed and non-toothed), curved grasping forceps, clip applicator and curved scissors to complete the procedure. A limitation of the technique was the inability to use electro-diathermy within an unventilated skills laboratory.

Each cadaveric porcine LC procedure was recorded using a previously validated, video-motion analysis system (ROVIMAS) which enables synchronized recording of video and dexterity data, as described in chapter four (Dosis *et al.*, 2005). Dexterity data are recorded by sensors placed on the dorsum of the surgeon's hands, which relay movements of the hands to a commercially available motion tracking device (Isotrack II, Polhemus, Colchester, VT, USA). Custom-made software calculates useful dexterity measures such as path length and number of movements of each hand, together with time taken. Qualitative analysis of the video was carried out by two experienced laparoscopic surgeons blinded to the identity and trial of each surgeon, marking procedures on the OSATS-based validated global rating scale (Martin *et al.*, 1997).

5.2.3 Proficiency data collection

In order to define benchmark levels of proficiency on the cadaveric porcine procedure, 10 experienced surgeons (>100 LCs each) were recruited to complete two cases each. These were conducted and assessed in the same manner as described for the novice subjects. The reason for including two cases was to ensure that all surgeons were familiarised to the set-up on their first run, and the results from the second run collated for definition of benchmark levels.

5.3 Statistical Analysis

The recruitment of 10 subjects per group was based upon power calculations from previous studies on virtual reality simulation. An inter-group analysis between the two novice groups for baseline laparoscopic skills and between successive laparoscopic cholecystectomies utilised the Mann-Whitney U Test, in accordance with the non-parametric nature of the data. Similarly statistical analysis was performed to enable comparisons between novice groups and the data of experienced surgeons. Inter-rater reliability for video scores was analysed with the Cronbach's alpha test statistic. The Statistical Package for the Social Sciences version 11.5 (SPSS, Chicago, Illinois, USA) was used in this analysis.

Examination of learning curves was with a longitudinal data analysis technique, based upon the correlation among repeated measurements in individual subjects, and variables that relate both to the average effect and the nature of variability. Specific to this study, the two variables of interest are change in operative performance over time, i.e. learning curves (intra-group progression), and differences in operative performance over time between control and VR-trained groups (inter-group progression). The repeated measures data was analyzed by using Generalized Estimating Equations (GEE) approach developed by Zeger et al. in the late 1908s (Liangand and Zeger, 1986;Zeger et al., 1988).

Marginal models (autoregressive matrix) based on generalised estimation equations were applied to perform regression analysis. The analysis took account of the fact that modelling the variation in surgical performance to perform tasks (i.e. learning curve) should be based primarily within subjects rather than between groups. For this reason the parameters of generalized estimating equations represent the average difference between subjects. In order to enable a more complex pattern of correlations among the repeated observations, an autoregressive structure was used, based upon the observation that correlations decrease substantially as the separation between observations increase. Learning curve analysis was conducted with Intercooled Stata version 8.0 for Windows (Stata Corporation, USA).

Finally, in order to calculate the transfer-effectiveness ratio (TER) for control versus VR-trained groups, the equation below was used:

$$TER = \frac{Yo - Yx}{X}$$

where Yo is the median time required by the control group to reach performance criterion and Yx is the corresponding measure for the VR trained group after having received a median of X amount of training time on the simulator, which is also the denominator (Taylor *et al.*, 2005).

5.4 Results

5.4.1 Baseline laparoscopic skill of novices

Baseline scores of laparoscopic skill on the 'glove model' did not demonstrate any significant differences between the control and VR-trained groups for parameters of time taken (median 562.5 vs. 555.6 seconds, p=0.971), path length (17.57 vs. 13.42 metres, p=0.529), number of movements (206 vs. 198, p=0.579) and error scores (23 vs. 22, p=0.398).

5.4.2 Completion of the virtual reality training curriculum

Of the 10 subjects randomised to the VR-trained group, 9 managed to complete the curriculum through attainment of pre-defined proficiency levels of performance. One subject did not manage to complete all VR sessions due to timetabling difficulties, and was thus disregarded from subsequent analyses. Of the nine who completed the VR curriculum, the median training time (*X*) as recorded by the simulator was 3967 seconds (range 3030 – 6738) over a median of 41 separate sessions (range 38 – 55).

5.4.3 Inter-group comparisons on porcine laparoscopic cholecystectomies

Inter-group comparisons revealed significant differences in performance on the first LC between control and VR trained groups for time taken (4590 vs. 2165 seconds, p=0.038), total path length (169 vs. 87 metres, p=0.001), total number of movements (2446 vs. 1029, p=0.009) and video rating scores (17 vs. 25 [out of 35], p=0.001). Though the VR-trained group performed only three LCs each, they consistently outperformed the control group on all measured parameters, as shown in tables 5a to 5e.

Statistical equivalence of performance was achieved between the fifth LC for the control group and the third LC for the VR-trained group on dexterity-based parameters: time taken (1598 vs. 1365 seconds, p=0.131), total path length (86 vs. 49 metres, p=0.110), total number of movements (875 vs. 647, p=0.110), though video rating scores (out of 35) remained significantly different in favour of the VR-trained group (25 vs. 31, p=0.003). Inter-rater reliability analysis revealed $\alpha = 0.74$.

Table 5a – Statistical analysis for first laparoscopic cholecystectomy

FIRST SESSION	*Control group (n=10)*	*VR trained group (n=9)*	*P (Control vs. VR)*	*Experienced surgeons (n=10)*	*P (Control vs. exp)*	*P (VR vs. exp)*
Total time (seconds)	4590 (3041 – 5044)	2165 (1898 – 3409)	0.038	765 (694 – 962)	<0.001	<0.001
Total path length (metres)	169 (136 – 308)	87 (64 – 112)	0.009	73 (47 – 99)	0.001	0.369
Total number of movements	2446 (1593 – 2838)	1029 (822 – 1565)	0.009	506 (388 – 682)	<0.001	0.001
Global rating scale (out of 35)	17 (8 – 27)	25 (11 – 33)	0.001	33 (29 – 35)	<0.001	<0.001

Table 5b – Statistical analysis for second laparoscopic cholecystectomy

SECOND SESSION	*Control group (n=10)*	*VR trained group (n=9)*	*P (Control vs. VR)*	*Experienced surgeons (n=10)*	*P (Control vs. exp)*	*P (VR vs. exp)*
Total time (seconds)	2908 (1567 – 3891)	2170 (1224 – 2739)	0.102	765 (694 – 962)	<0.001	0.001
Total path length (metres)	145 (95 – 253)	81 (50 – 97)	0.034	73 (47 – 99)	0.016	0.870
Total number of movements	1699 (774 – 2224)	1014 (583 – 1252)	0.034	506 (388 – 682)	0.001	0.006
Global rating scale (out of 35)	21 (14 – 30)	30 (24 – 34)	0.002	33 (29 – 35)	<0.001	0.133

Table 5c – Statistical analysis for third laparoscopic cholecystectomy

THIRD SESSION	*Control group (n=10)*	*VR trained group (n=9)*	*P (Control vs. VR)*	*Experienced surgeons (n=10)*	*P (Control vs. exp)*	*P (VR vs. exp)*
Total time (seconds)	2740 (1870 – 3875)	1365 (1184 – 1524)	0.016	765 (694 – 962)	<0.001	0.006
Total path length (metres)	145 (64 – 272)	49 (40 – 79)	0.005	73 (47 – 99)	0.102	0.274
Total number of movements	1631 (862 – 2054)	647 (590 – 796)	0.005	506 (388 – 682)	0.001	0.055
Global rating scale (out of 35)	24 (15 – 34)	31 (21 – 35)	0.001	33 (29 – 35)	<0.001	0.236

Table 5d – Statistical analysis for fourth laparoscopic cholecystectomy

FOURTH SESSION (control group only)	*Control group (n=10)*	*VR trained group (n=9)*	*P (Control vs. VR)*	*Experienced surgeons (n=10)*	*P (Control vs. exp)*
Total time (seconds)	2923 (1563 – 3572)	1365 (1184 – 1524)	0.010	765 (694 – 962)	0.001
Total path length (metres)	146 (53 – 224)	49 (40 – 79)	0.016	73 (47 – 99)	0.049
Total number of movements	1748 (814 – 1897)	647 (590 – 796)	0.016	506 (388 – 682)	0.002
Global rating scale (out of 35)	23 (12 – 31)	31 (21 – 35)	<0.001	33 (29 – 35)	<0.001

Table 5e – Statistical analysis for fifth laparoscopic cholecystectomy

FIFTH SESSION (control group only)	*Control group* *(n=10)*	*VR trained group* *(n=9)*	*P* *(Control vs. VR)*	*Experienced surgeons (n=10)*	*P* *(Control vs. exp)*
Total time (seconds)	1598 (1394 – 2396)	1365 (1184 – 1524)	0.131	765 (694 – 962)	0.001
Total path length (metres)	86 (47 – 179)	49 (40 – 79)	0.110	73 (47 – 99)	0.436
Total number of movements	875 (672 – 1336)	647 (590 – 796)	0.110	506 (388 – 682)	0.002
Global rating scale (out of 25)	25 (18 – 29)	31 (21 – 35)	0.003	33 (29 – 35)	<0.001

5.4.4 Learning curve analysis

Learning curve analysis for repeated sessions, and between VR and control groups, revealed statistically significant learning curves for both the control and VR-trained groups in terms of time taken, total path length, number of movements (table 5f, $p<0.001$ for all). The learning curves for both control and VR-trained groups on cadaveric porcine LCs are illustrated in figures 5c to 5f.

Table 5f – Results of GEE analysis-effect of virtual reality training in surgical performance parameters

	B-COEFFICIENT	STANDARD ERROR	p-VALUE	95% CI
Total Time (seconds)				
Session	-519.7	87.5	<0.001	-691.2 to -348.2
Virtual Learning	-1741.4	306.8	<0.001	-2342 to -1139.9
Constant	4329	340.6	<0.001	3661.4 to 4996.6
Total number of movements				
Session	-361.4	60.8	<0.001	-480.6 to -242.2
Virtual Learning	-1459	216.2	<0.001	-1882.8 to -1035.1
Constant	2551.5	238.11	<0.001	2084.8 to 3018.2
Total path length (metres)				
Session	-134.6	27.9	<0.001	-189.3 to -79.9
Virtual Learning	-566.4	95.7	<0.001	-754 to -378.8
Constant	517	107.6	<0.001	306 to 728

Figure 5c – Learning curves for control and VR-trained groups for time taken

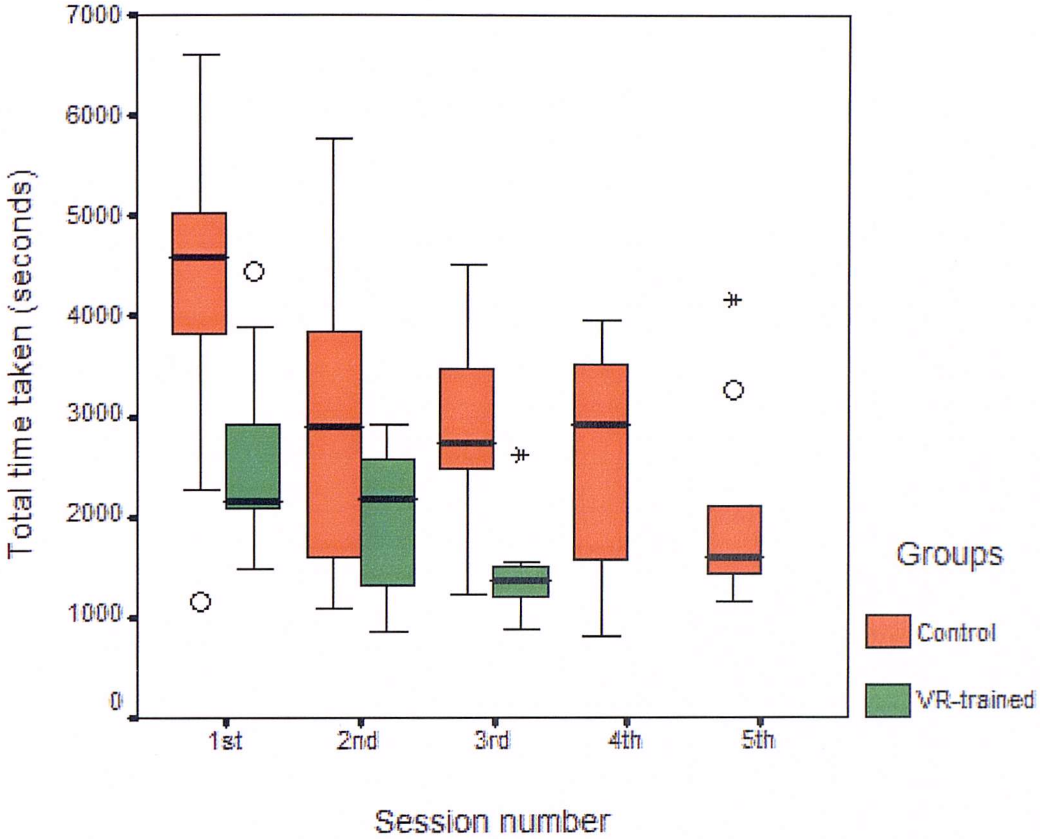


Figure 5d – Learning curves for control and VR-trained groups for total path length

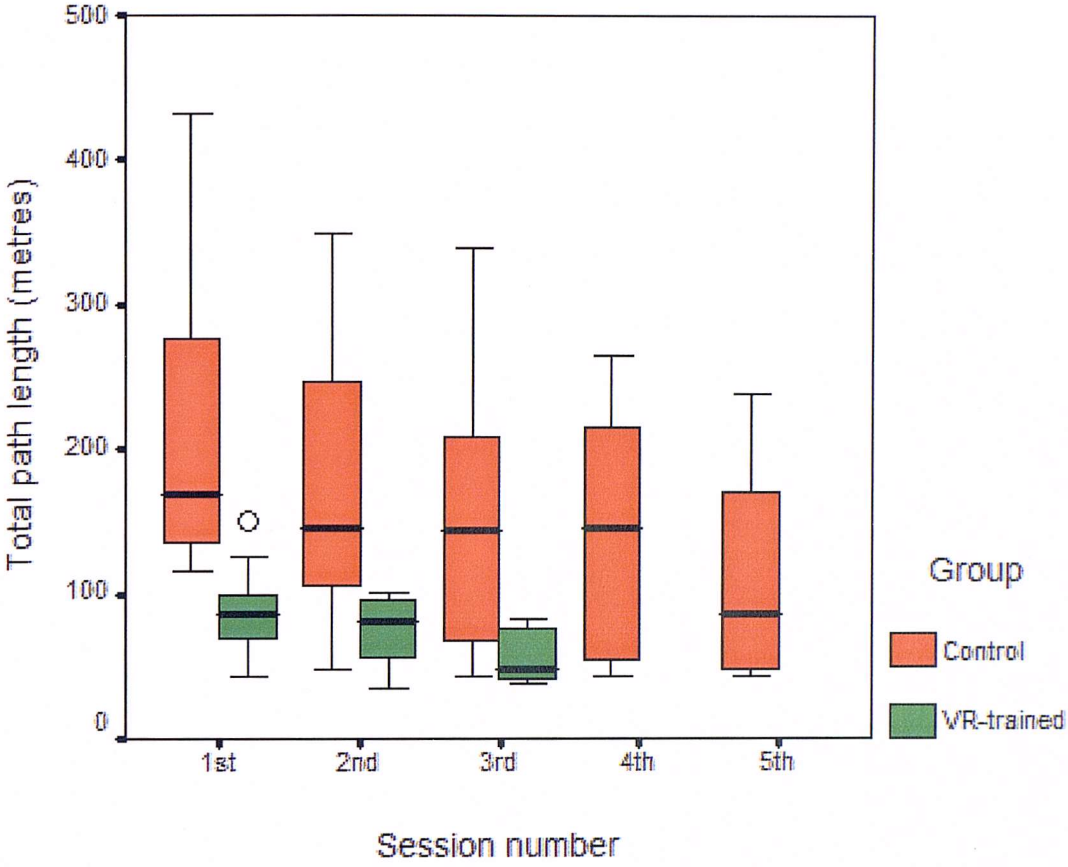


Figure 5e – Learning curves for control and VR-trained groups for total number of movements

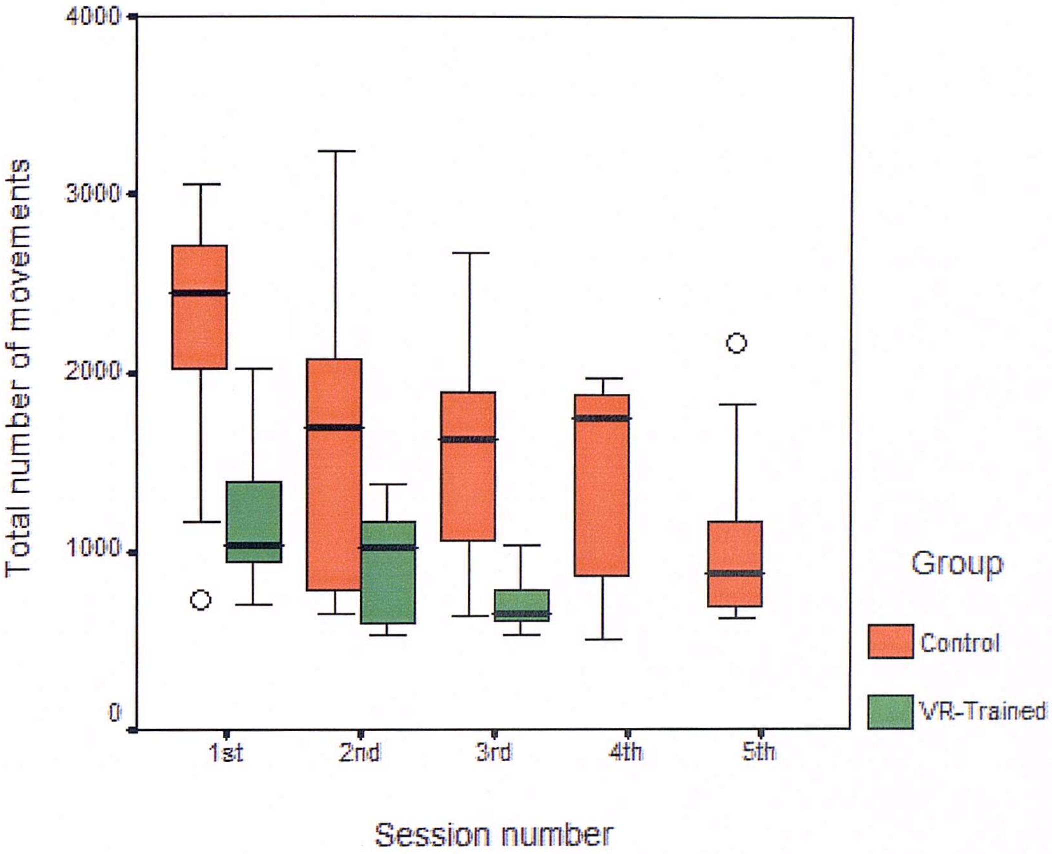
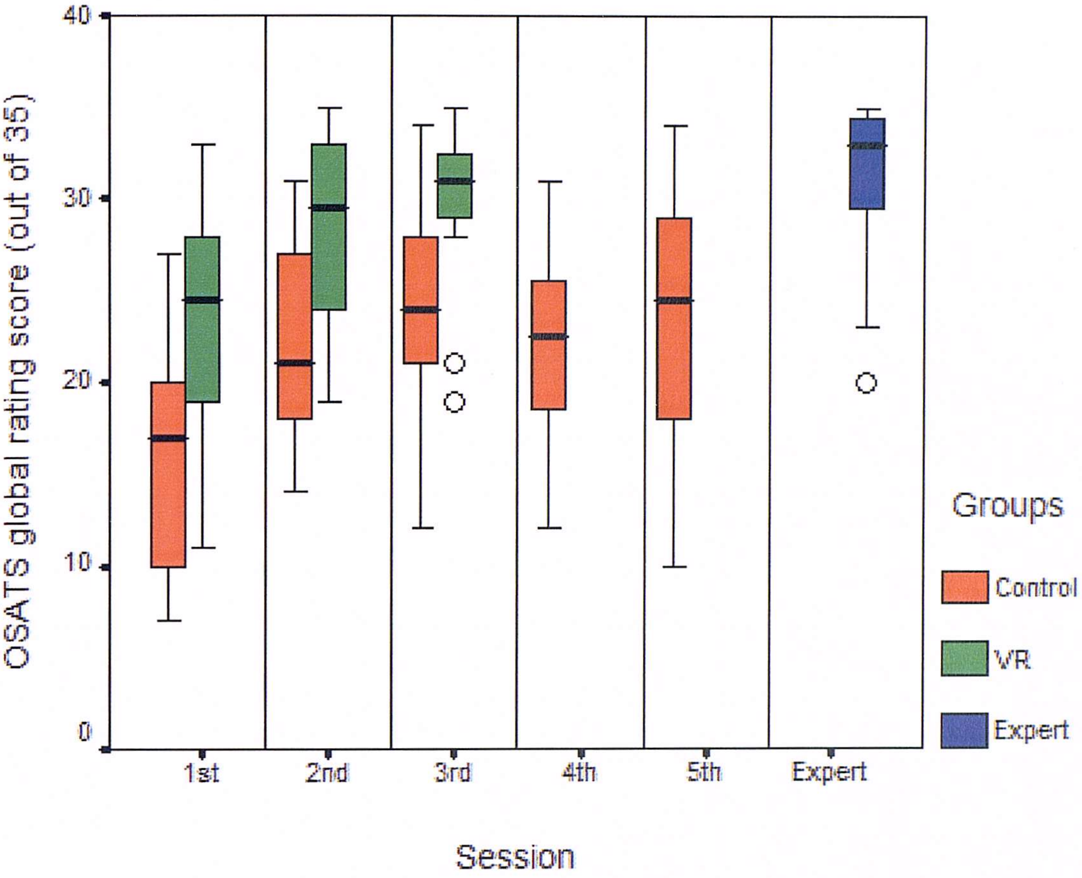


Figure 5f – Learning curves for control and VR-trained groups for video-based rating score, and benchmark levels of experienced surgeons



5.4.5 Comparison of performance to expert levels

In comparison to the experienced group, both control and VR-trained groups were significantly slower and made greater numbers of movements to complete the first LC (table 5a). However, the VR-trained group achieved statistically equivalent path lengths to the experienced group for this first session (87 vs. 73, p=0.369). By their third session, the VR-trained group had also achieved statistically equivalent scores to the experienced group for total number of movements (647 vs. 546, p=0.055) and video rating scores (31 vs. 33, p=0.236), though not for time taken (1365 vs. 765, p=0.006) (table 5c). With regard to the control group at their fifth LC, only total path length was equivalent to that of the experienced surgeons (86 vs. 73, p=0.436) (table 5e).

5.4.6 Calculation of the transfer-effectiveness ratio

Though the aim was to calculate the transfer-effectiveness ratio from time taken to achieve proficiency-based criteria (i.e. benchmark scores from the experienced group), this was not possible as neither group had achieved all the pre-defined criteria by their last session. It is however possible to equate equivalence of training by making use of the fact that equivalence of performance on dexterity parameters was achieved between the control and VR-trained groups at their fifth and third sessions respectively.

The control group required a median of 14,759 seconds (Yo) to complete five cadaveric porcine LCs, whilst the VR-trained group required a median of 5,700 seconds (Yx) to complete three LCs. In addition, the VR-trained group underwent a median of 3967 seconds (X) training time on the simulator. Applying this data to the TER equation results in a calculation of:

$$TER = \frac{14759 \text{ seconds} - 5700 \text{ seconds}}{3967 \text{ seconds}} = 2.28$$

5.5. Conclusions

The training of surgeons is a subject of broad concern to health professionals, patients, government officials and the public alike (Jackson B, 1999;Debas et al., 2005). Reports of medical error within the healthcare and public domains have driven the need to define objective and valid measures of competence prior to credentialing of surgeons for independent practice (Darzi et al., 1999;Aggarwal and Darzi, 2006). This is tempered by the implementation of reduced working hours for junior doctors on both sides of the Atlantic. The medical community is thus obliged to develop and maintain new training paradigms which can deliver competent practitioners without undue harm to patients during the acquisition of these skills. Since the late 1970s, two to three day surgical skills workshops have taught junior surgeons basic techniques on synthetic and animal-based tissues. More recently, the development of virtual reality systems have been at the forefront of training in minimally invasive techniques such as laparoscopy, endoscopy and bronchoscopy (Aggarwal *et al.*, 2004b). Though studies have demonstrated the ability of training on a simulated model to improve real performance on patients, there remains a lack of integration of simulation-based training into daily practice.

The aim of this study was firstly to determine the transfer of skill of a proficiency-based virtual reality training curriculum to real procedures. The VR-trained group performed significantly better at the first laparoscopic cholecystectomy than the control group. Nonetheless, both groups demonstrated statistically significant learning curves over subsequent procedures, demonstrating the fact that VR-based training is not a substitute, but rather an adjunct to traditional modes of training.

The second aim was to determine the hypothesized differences between learning curves on real cases for VR-trained and control groups of novice laparoscopic surgeons, and how this related to time spent on the VR simulator, i.e. the *transfer effectiveness-ratio*, or TER of a VR-based surgical training curriculum. Similar studies in the aviation industry have described a TER of 0.5, though the results of this study have defined a TER of 2.28, i.e. every minute that was spent on the VR simulator was equivalent to 2.28 minutes on the cadaveric porcine cholecystectomy.

In order to qualify this result, it is important to mention that training on the VR curriculum was stepwise and proficiency-based. This ensured that the VR-trained subjects acquired skills in a hierarchical and task-based manner and did not progress onto a more difficult task until they had achieved benchmark levels of skill on the preceding task (Fitts AM and Posner MI, 1967). Thus it was not just hands-on practice on the simulator which led to improved performance on the porcine LCs, but the step-wise and proficiency-based nature of the simulator-based training curriculum. This was despite the fact that both groups acquired cognitive knowledge, confirming the importance of hands-on practice in a simulated setting (Immenroth *et al.*, 2007).

A further advantage of this study is the use of novel methodology within the surgical skills literature to model the variation in two levels, i.e. both within and between groups (Liangand and Zeger, 1986;Zeger et al., 1988). It is also taken into account that surgical performance during repeated tasks should be modeled on a subject rather than group basis. Though the group as a whole may be improving their performance, it is more informative to know the within subject variation over successive procedures. The GEE approach accounts for this with the assumption that over successive repetitions, within-group variation should decrease and between-group variation increase.

With regard to the VR simulation device, it may be considered to be of moderate fidelity (or realism) in that the software is of high quality, but the hardware does not incorporate force feedback (Lamata *et al.*, 2006). Though VR-trained subjects acquired skills on the cadaveric porcine model at a faster rate, leading to a shorter and flatter learning curve, than the control group, it must be noted that it is not the simulator but rather the mode of simulation-based training (i.e. proficiency-based, stepwise and structured) which has led to this result.

One of the major problems today is the lack of funds to purchase and maintain surgical simulators. The next stage of research would be to calculate the TER for other surgical simulator-based curricula, enabling a marker of cost-effectiveness to be developed in relation to initial outlay costs. With further studies such as this one, it may be possible to define the time and cost reductions for training on specific VR simulation-based curricula compared to traditional modes of acquiring surgical skill, for example, a notion of *'number needed to train'* to achieve benchmark criteria.

It must also be noted that the aim of a TER is to define how long it takes for the subjects to achieve proficiency. This was one of the drawbacks of this study, as time and financial limitations meant that it was only possible to perform five LCs for the control group and three LCs for the VR-trained group. In fact, the VR-trained group achieved proficiency on the cadaveric LCs for three out of four measures, but fell short on time taken. The control group only achieved expert benchmark levels on the parameter of total path length, despite five LC procedures.

However, the control group had to perform complete porcine LCs with only minimal knowledge from the one day seminar, i.e. their training session were not structured in a step-wise manner. This does not happen during patient cases; the trainee more often spends time holding the laparoscope, followed by performing parts of each procedure, until eventually they can perform one whole procedure from 'skin to skin'. This is the nature of the apprenticeship model, and extends the learning curve over tens to hundreds of real cases (Halsted, 1904). Whilst a stepwise approach enables the trainee to practice their skills on patients, there is also the propensity to lead to unnecessary complications. Furthermore, the operating room is not an educationally orientated environment, training occurs by chance, and can lead to increased stress which may have a negative effect upon the learning process. Simulation-based training has the benefit of putting the trainee's goals first, enabling one to make mistakes and most importantly learn from them through constructive feedback delivered by the simulator and or trainer.

A further limitation of this study is the use of cadaveric porcine LCs placed in a video-box trainer as opposed to real patient cases. It could be argued that porcine LCs are easier than patient cases, and thus the results of this study are merely of interest rather than definitive. Though a valid criticism, it must also be considered that the majority of training courses in basic laparoscopic skills utilize such models prior to enabling young surgeons to progress onto performing real cases. Furthermore, the differences in performance between the control group and the experienced surgeons at first and second sessions confirm the construct validity of the porcine model to assess laparoscopic surgical skill at performing a human LC, i.e. performance on the porcine model is related to real surgical experience, and thus the model measures what it purports to measure (Moorthy *et al.*, 2003). It is thus acceptable to relate the results of this study to real patient cases. Ultimately, it will be important to perform such a study not only to define the TER of other VR surgical simulation curricula, but also of traditional video-box training curricula as utilized during laparoscopic skills training courses, and their transfer to patient cases.

The aim of this study has been to prove not only that training on a VR-based laparoscopic surgical curriculum improves performance on real cases, but also that this effect persists beyond the first procedure. The definitive aim of simulation-based training curricula in any domain is to reduce the time or number of cases taken to achieve proficiency within the real environment in order to reduce costs, time and errors within that real environment (Gallagher and Cates, 2004). It is with this concept that we hope to challenge current methods of training by deploying adjunctive environments such as surgical skills laboratories to train the surgeons of the future (Aggarwal and Darzi, 2006). A further aim of future research within this field should also be to define the proficiency levels of factors other than technical skills which lead to the development of surgical competence. Within an operative environment, communication, team working, decision making and judgment are key domains which must also be honed to ensure one possesses the ability for independent and competent practice (Aggarwal et al., 2004c;Moorthy et al., 2005;Moorthy et al., 2006).

Chapter 6

TRAINING IN LAPAROSCOPIC ROUX-EN-Y GASTRIC BYPASS – AN EVIDENCE-BASED APPROACH

Aim

Construct an evidence-based training program for an advanced laparoscopic surgical procedure, which is task and proficiency-based, though based upon cadaveric porcine models for assessment of technical skill.

6.1 Introduction

Morbid obesity is a worldwide problem, and its incidence is increasing at an alarming rate (Yanovski and Yanovski, 2002). Surgical therapy has been shown to result in significant and sustainable weight loss in comparison to medical management of the disease (Colquitt et al., 2003;Sjostrom et al., 2004;Maggard et al., 2005). A number of prospective randomised trials have declared Roux-en-Y gastric bypass to be the surgical procedure of choice (Sugerman et al., 1987;Hall et al., 1990;Weber et al., 2004;Olbers et al., 2005). The operation, first described by Mason and Ito in 1967, involves creation of a 20-30ml gastric pouch (the restrictive element), attached to a Roux loop of jejunum to decrease fat absorption (the malabsorptive element) (Mason and Ito, 1967). However, operating on this group of patients is extremely challenging, and can lead to a number of post-operative complications. In 1993, Wittgrove and Clark performed the first laparoscopic Roux-en-Y gastric by-pass (LRYGBP) procedure (Wittgrove *et al.*, 1994). The patients no longer had to tolerate a large midline laparotomy, and its associated complications, confirmed by results from the series of Nguyen et al. consistently reporting superior outcomes with the minimally invasive approach (Nguyen et al., 2000;Nguyen et al., 2001a;Nguyen et al., 2001b;Nguyen and Wolfe, 2002).

The challenges posed by operating on an obese abdomen with the open approach had been minimised, though newer advanced skills were needed to acquire proficiency in LRYGBP. The learning curve for this procedure has been quoted to be between 75 and 100 cases, during which time complication and conversion rates may be twice as high as those of experienced surgeons (Schauer et al., 2003;Nguyen et al., 2003;Oliak et al., 2003;Ballesta-Lopez et al., 2005;Ballantyne et al., 2005;Shikora et al., 2005). The reason for this long and arduous period of training is due to the complexity of both the operative procedure, and the skills necessary to carry out the procedure. The surgeon must work in more than one abdominal quadrant and perform tasks such as gastric pouch creation and Roux limb formation. Advanced skills for bowel manipulation, laparoscopic suturing and the use of new tools (laparoscopic staplers, high energy devices) must also be mastered (Higa *et al.*, 2000).

In accordance with the apprenticeship model of surgical training, the majority of these skills are learnt through practice on patients. Over 100 minimally invasive fellowships are available each year in the United States, enabling post-residency surgeons to perform cases to scale the learning curve (2005). Entry to these posts is highly competitive, especially at high volume centres. Furthermore, many surgeons enter these training programs with minimal proficiency beyond basic laparoscopy (i.e. diagnostic, cholecystectomy and appendectomy) (Kothari *et al.*, 2005). It is not unusual for a surgeon to spend many minutes to place a single intracorporeal suture at commencement of their learning curve, leading to procedures which can last for many hours. Furthermore, it has been shown that these skills can be learnt outside the operating room (Aggarwal *et al.*, 2006b).

There has recently been a drive by The Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) to encourage the commencement of advanced laparoscopic training during residency (1998;Rattner et al., 2001;2003). The proposals prescribe the role of skills laboratories (using animate and inanimate material), courses and mini-fellowships, followed by preceptorship at accredited centres., i.e. an attempt at the development of a training curriculum The importance of objective assessment of performance is also stated, verifying the fact that duration of training to achieve proficiency may vary between individuals. The underlying principle is to augment training prior to practice on patients, with a concomitant reduction of the learning curve in the operating room. It is with this approach, in alignment to the methodology presented in chapter 2, that it was sought to define an evidence-based curriculum for commencement of training in the advanced laparoscopic procedure of LRYGBP.

6.2 Methods

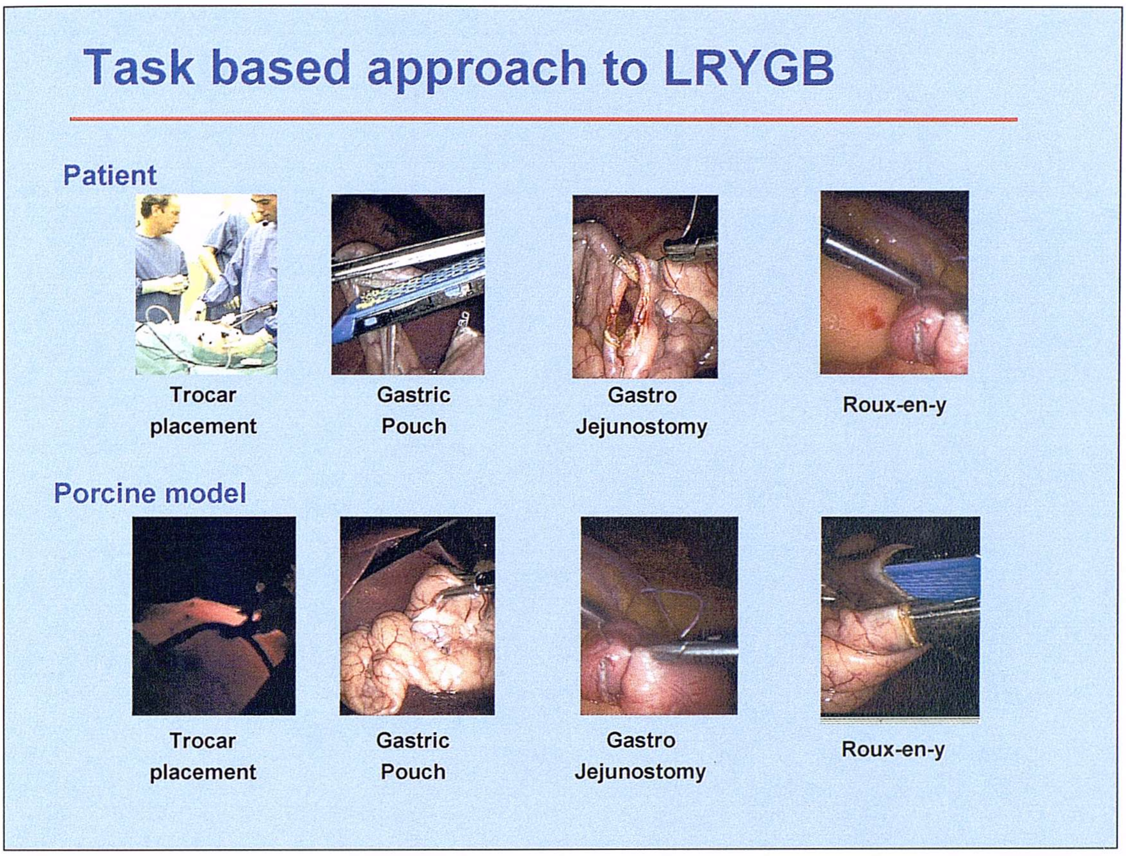
6.2.1 Development of a task-based approach to LRYGBP

Learning a complex procedure such as LRYGBP is difficult. In the fields of sport and music, acquisition of skills occurs through stepwise progression, followed by merging of the individual blocks into a smooth and controlled action (Ericsson, 2004). To some extent, this occurs contemporaneously during surgical procedures. The approach was to emphasise the division of LRYGBP into four key tasks: trocar placement, jejuno-jejunostomy, creation of gastric pouch and gastro-jejunostomy. The aim was for surgeons new to the operation to be able to digest individual tasks prior to tackling the entire procedure.

6.2.1.1 A multimedia-based training aid

As an aid to learning the procedure, a Digital Versatile Disc (DVD) was produced to prescribe the task based approach (figure 6a). The emphasis was upon a step-by-step process of each individual task, for both the patient and porcine models. The aim was not only to develop a technical skills learning tool, but also to define the limitations of the porcine model when compared to the human procedure. The DVD was subsequently incorporated into a two-day morbid obesity master-class which entailed a total of six hours of hands-on practice on anaesthetised porcine models. At each operating table, a second screen was available to display the DVD, enabling surgeons to mimic each step of the procedure.

Figure 6a Start-up screen on the Digital Versatile Disc for Training in LRYGB



Task based approach to LRYGB

Patient

Trocar placement
Gastric Pouch
Gastro Jejunostomy
Roux-en-y

Porcine model

Trocar placement
Gastric Pouch
Gastro Jejunostomy
Roux-en-y

6.2.2 An assessment tool for LRYGBP

Within the task-based approach to LRYGBP, it is necessary not only to learn the procedure in definitive stages, but also to objectively assess the acquisition of skill. Any assessment tool must be feasible, objective, valid, reliable and cheap to deliver (Moorthy *et al.*, 2003). Currently, no assessment tool exists to perform this function for LRYGBP.

6.2.2.1 Definition of an assessment task

The aim was to produce a high-fidelity model which did not require specialist handling and storage facilities. Furthermore, instead of producing an assessment task for each individual step, we attempted to incorporate the individual skills of bowel manipulation, intracorporeal suturing and use of staplers into one task. To aid this, two experienced bariatric surgeons reviewed videos of experienced and inexperienced surgeons performing LRYGBP on patients. The sentinel task for assessment of skill performed by surgeons in a standardised manner was defined to be the side-to-side stapled jejuno-jejunostomy.

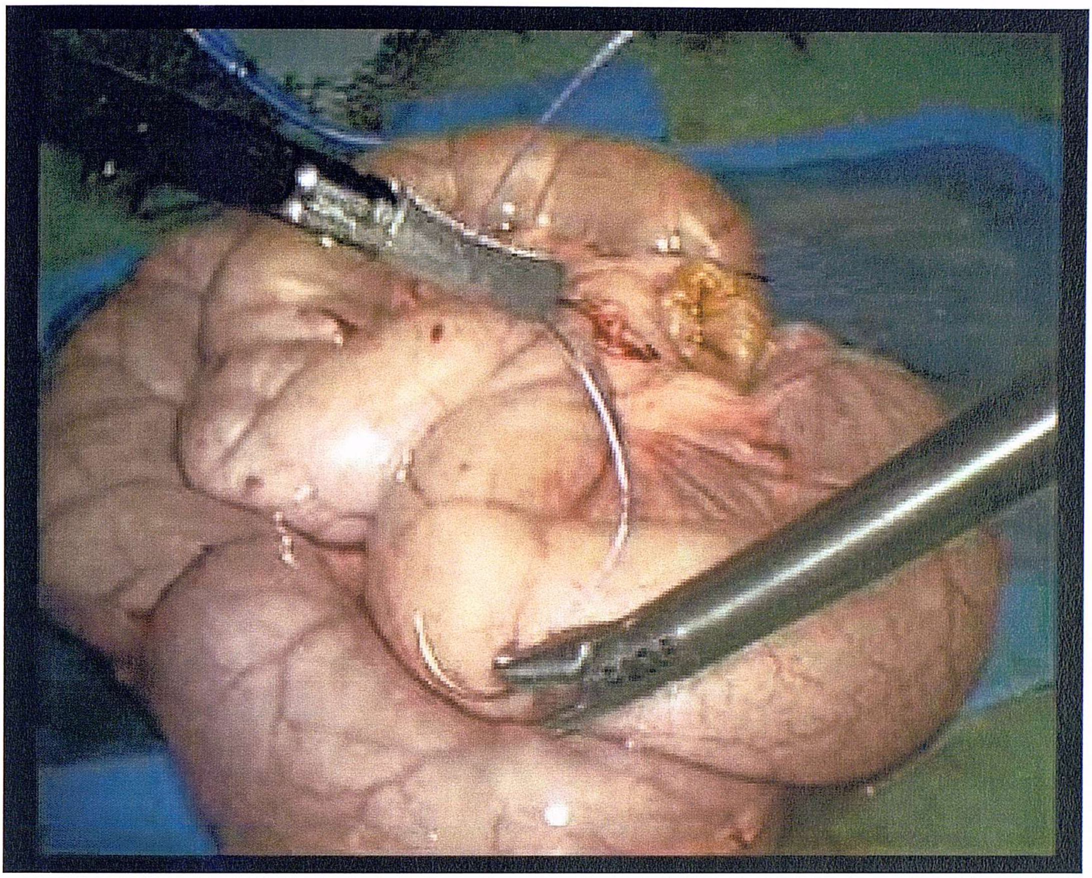
6.2.2.2 Development of an assessment model

To develop a cheap and feasible model for this part of the operation, cadaveric porcine material was placed into a video-box trainer. However, manipulation and positioning is difficult due to collapse of the bowel from lack of gastrointestinal contents. The addition of a thickened solution normally used to treat dysphagia (Thick and Easy, Hormel Foods, Austin, MN, USA) restored a realistic appearance to the bowel (figure 6b). Initially, two pieces of bowel, each 30cm in length were placed into a video-box trainer onto a cork board, and fixed to the cork by the mesentery with a heavy-duty office stapler. This enabled subjects to manipulate the two bowel segments as they would prior to performing the jejuno-jejunostomy. The jejuno-jejunostomy was performed by placing an intracorporeally tied stay suture, then making enterotomies in each bowel limb with scissors, diathermy (if the bowel was placed onto a diathermy pad) or harmonic scalpel, followed by firing of a linear stapler between the two bowel lengths, and finally closure of the enterotomy with a running suture. The model was subsequently modified to use 50cm lengths of bowel which were placed in a U-configuration, providing the illusion of two side-by-side pieces of bowel.

6.2.2.3 Face and content validity of the assessment model

The cadaveric porcine model was easy to produce, did not require any specialist storage or handling facilities, and cheap to use. With regard to validity, it possessed both *face* (whether the model resembles the task it is based upon) and *content* validity (the extent to which the model measures surgical skill and not simply anatomical knowledge). In order to use this model for assessment of technical skill, it was necessary to confirm the *construct* validity of the task, i.e. does it measure the traits one purports measure. This can be inferred by measuring differences in performance between experienced and inexperienced subjects, in an objective manner(Moorthy *et al.*, 2003).

Figure 6b Intra-operative view of the cadaveric porcine jejuno-jejunostomy model



6.2.2.4 Construct validity of the cadaveric porcine model

The assessment of a surgeon's technical skill for an advanced laparoscopic task necessitates baseline laparoscopic experience. Thus, inclusion criteria for this study were completion of >100 basic and/or intermediate laparoscopic procedures (table 6a).

Subsequently, 27 surgeons were recruited to perform the laparoscopic stapled jejuno-jejunostomy in a video-box trainer. They were divided according to experience in advanced laparoscopic procedures (table 6a) into 11 inexperienced (<10 advanced procedures), eight intermediate (20-50 advanced procedures) and eight experienced (>100 advanced procedures) surgeons.

At commencement, all subjects were familiarized with the tools and shown a short video clip of the task to be performed. They then performed the task with hook diathermy for enterotomy, and a linear laparoscopic stapler (EndoGIA 45mm stapler, U.S. Surgical, Norwalk, CT, USA) for creation of the jejuno-jejunostomy. Intracorporeal suturing was performed with the use a laparoscopic needle holder (Richard Wolf Medical Instruments, Vernon Hills, IL, USA).

Table 6a Definition of basic, intermediate and advanced laparoscopic procedures

Basic Laparoscopic Procedure	***Advanced Laparoscopic Procedure***
Diagnostic Laparoscopy Cholecystectomy Appendectomy	Bariatric procedures (excluding banding) Colorectal procedures Oesophagectomy Gastrectomy Pancreatectomy Splenectomy Adrenalectomy
Intermediate Laparoscopic Procedure	
Nissen fundoplication Ventral hernia repair	

Objective assessment of performance was by motion analysis for dexterity, and video analysis for quality of the procedure performed. As before, motion parameters were recorded by the Robotic Video Motion Analysis Software (ROVIMAS) which involves the placement of sensors on the back of a surgeon's hands (Dosis *et al.*, 2005). A commercially available device (Isotrack II^{TM}, Polhemus, VT, USA) emits electromagnetic waves to track the position of the sensors in *x*, *y* and *z* axes 20 times per second. The device runs from a standard laptop computer and bespoke software enables the data to be analysed for time taken, total distance traveled (i.e. path length) and total number of movements by each hand.

The ROVIMAS system was also used to synchronously record the video of the procedure, direct from the laparoscopic stack. All videos were analysed post-hoc using a validated global rating scale, by two surgeons experienced in bariatric surgical techniques, in a blinded manner (Martin *et al.*, 1997). Five categories from the previously validated scale were used to mark technical skill, each rated on a Likert scale from 1 to 5 (table 6b), with a maximum possible score of 25.

There was some concern that the generic nature of this scale would not be able to capture all of the elements of this task. A procedure-specific rating scale was thus developed, with four categories, each marked from 1 to 5, with a possible total of 20 (table 6c).

Table 6b Modified generic global rating scale of operative skill

	1	2	3	4	5
Respect For Tissue	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal damage
Time & Motion	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Economy of movement and maximum efficiency
Instrument Handling	Repeatedly makes tentative or awkward moves with *instruments*		Competent use of instruments although occasionally appeared stiff or awkward		Fluid moves with instruments and no awkwardness
Flow of Operation & Forward Planning	Frequently stopped operating or needed to discuss next move		Demonstrated ability for forward planning with steady progression of operative procedure		Obviously planned course of operation with effortless flow from one move to the next
Knowledge of Specific Procedure	Deficient knowledge. Needed specific instruction at most operative steps		Knew all important aspects of the operation		Demonstrated familiarity with all aspects of the operation

Table 6c Procedure-specific rating scale for skill in LRYGB

	1	2	3	4	5
Laparoscopic suture placement (stay)	Lack of dexterity in positioning needle, and driving through tissue. Does not attend to recognized knot tying techniques		Needle held in appropriate position; appropriate technique of knot tying, though fumbles occasionally		Accurate needle positioning, placement and smooth knot tying technique
Enterotomy	Placed in a haphazard manner. Poor relation between grasper and hook; excessively large or small		Appropriate size of enterotomy, though performed with some hesitation.		Appropriately sized and placed enterotomies, with no extra movements. Good relation of grasper and hook
Stapling	Unclear of how to use staple device. Drives staple jaws blindly into jejunum; closes jaws without both in bowel lumina		Uses staple device with hesitation. Uses stay suture to place jaws, though lacks appreciation of the ideal angle for insertion		Places staple jaws with ease, and uses stay suture to draw bowel into jaws. Smooth, controlled fire with no widening of enterotomies
Enterotomy closure	Poorly positioned stitch. Blindly placed continuous sutures with little regard to ensure enterotomy closure		Adequate stitch position. Sutures placed at varying distances apart, with gathering of bowel edges		Full thickness sutures placed at uniform distance apart

6.2.3 Application of the model as an assessment tool

Application of a research study to real practice is necessary to ensure the feasibility of an assessment tool in daily practice. It was thus decided to measure the change in technical skill following a period of training in LRYGBP. A pre- and post-course assessment of performance during the redesigned two-day laparoscopic morbid obesity master-class was carried out. The course involved lectures focused on surgical technique, hands-on practice on anaesthetized porcine models within a task-based approach, and faculty with extensive experience in LRYGBP.

All surgeons attending the two-day course were invited to participate, and 20 of the 22 attendees provided informed consent to be enrolled into the study. However, time constraints resulted in only 16 complete pre- and post-training data sets being available for analysis. The subjects performed a laparoscopic jejuno-jejunostomy on the cadaveric porcine model in a box trainer at commencement and conclusion of the course. In exactly the same manner as for the validation study, subjects were initially familiarized to the tools and shown a short video of the task. Assessment was by motion and video analysis, and performed for the whole task. In order to perform an additional assessment solely of intracorporeal suturing skills, data was extracted separately for initial stay suture placement with both motion analysis and marking of videos on a previously validated 29-point checklist for intracorporeal suturing (Moorthy *et al.*, 2004).

6.3 Statistical Analysis

Data was analysed with the Statistical Package for the Social Sciences version 11.5 (SPSS, Chicago, IL, USA) using non-parametric tests. Inter-group comparisons were analysed using the Kruskal-Wallis test for multiple groups and the Mann-Whitney test if only two groups were being compared. Intra-group comparisons between pre- and post-test performance were analysed with the Wilcoxon signed rank test. Inter-rater reliability of video scores was evaluated with Cronbach's alpha. A level of $p<0.05$ was considered statistically significant.

6.4 Results

6.4.1 Development of a task-based approach with an assessment tool for LRYGBP

The feedback from study participants, though not formally collated or evaluated, regarding the task-based approach to the course was positive, as were the face and content validities of the simulated jejuno-jejunostomy.

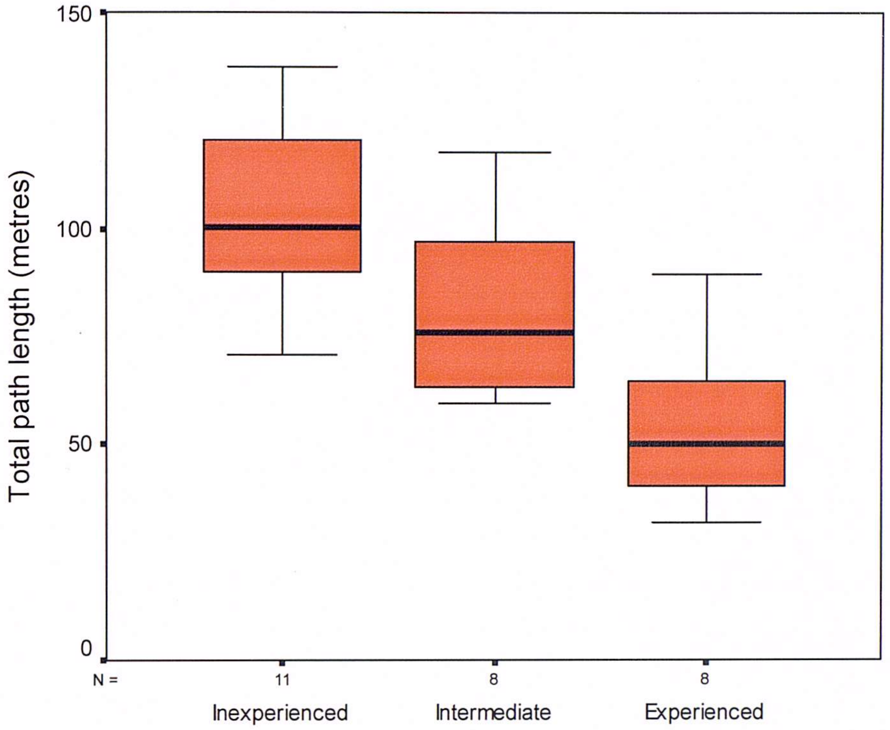
6.4.2 Construct validity of the cadaveric porcine model

A comparison across the three groups revealed a stepwise improvement in task performance from inexperienced to experienced subjects, confirmed by significant differences between all groups for each parameter (table 6d and figure 6c). The inter-rater reliabilities between the two raters for the generic and procedure-specific rating scales were $\alpha = 0.86$ and $\alpha = 0.76$ respectively.

Table 6d Construct validity of the jejuno-jejunostomy assessment tool; data is reported as medians, with interquartile range in parentheses

	Inexperienced *(n=11)*	*Intermediate* *(n=8)*	*Experienced* *(n=8)*	*P*
Total time (seconds)	1047 (880 – 1226)	823 (637 – 943)	531 (345 – 606)	<0.001
Total path length (metres)	100 (88 – 134)	76 (63 – 101)	50 (40 – 71)	0.001
Total number of movements	599 (512 – 675)	432 (358 – 543)	301 (211 – 369)	<0.001
Global rating scale (out of 25)	9 (7 – 14)	15 (12 – 20)	21 (17 – 24)	<0.001
Procedure-specific rating scale (out of 20)	8 (7 – 11)	13 (10 – 16)	17 (14 – 19)	<0.001

Figure 6c Box and whisker plot to illustrate construct validity for total path length between the three groups of surgeons (p=0.001)



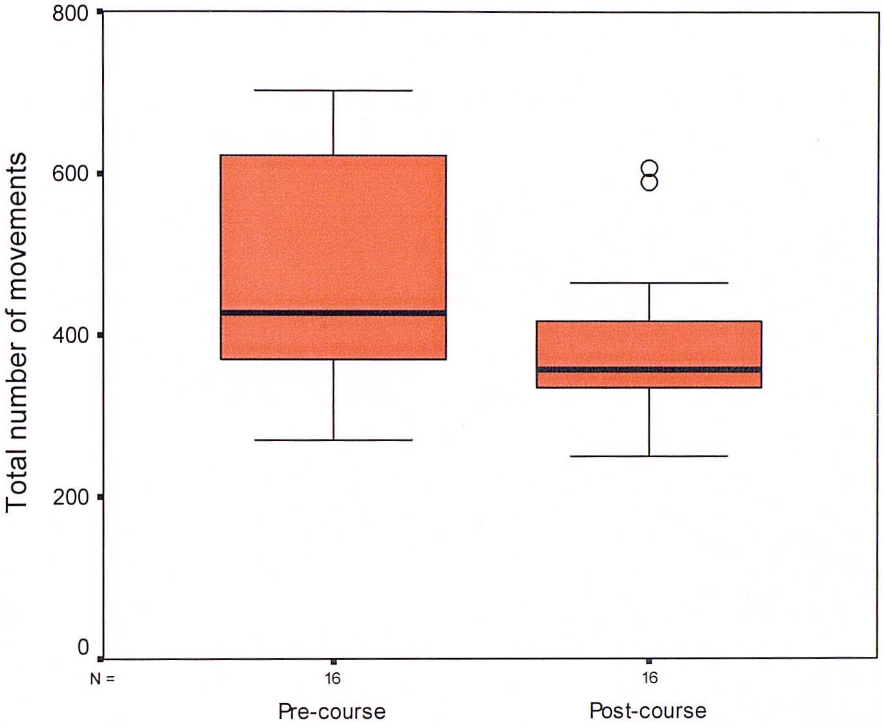
6.4.3 Application of the model as an assessment tool

There were significant improvements in performance on the model between the pre- and post-course assessments for motion tracking parameters though not for video-based rating scales (table 6e and figure 6d). When the performance of the 16 surgeons attending the course was compared with data from the eight experienced surgeons, there were significant differences at pre-course evaluations (table 6e). Apart from time taken, end of course parameters remained significantly different between experienced subjects and course participants.

Table 6e Comparison of performance at the whole procedure during pre- and post-course assessments, and with experienced surgeons; data is reported as medians, with interquartile range in parentheses

	Pre-course (n=16)	*Post-course (n=16)*	***P*** *(pre vs. post)*	*Experienced surgeons (n=8)*	***P*** *(pre vs. exp)*	***P*** *(post vs. exp)*
Total time (seconds)	753 (635 – 1055)	585 (534 – 795)	0.002	531 (345 – 606)	<0.001	0.120
Total path length (metres)	76 (63 – 107)	61 (57 – 80)	0.006	50 (40 – 71)	0.009	0.032
Total number of movements	427 (363 – 633)	357 (334 – 419)	0.003	301 (211 – 369)	0.005	0.038
Global rating scale (out of 25)	10 (8 – 14)	12 (9 – 14)	0.243	21 (17 – 24)	<0.001	<0.001
Procedure-specific rating scale (out of 20)	12 (8 – 14)	12 (9 – 14)	0.656	17 (14 – 19)	<0.001	<0.001

Figure 6d Box and whisker plot to illustrate pre- and post-course difference for total number of movements (p=0.003)



Assessment of performance for placement of the stay suture did not reveal significant improvements between pre- and post-course tests (table 6f). Furthermore, subjects were significantly worse than experienced surgeons at placing the stay suture, during both the pre- and post-course testing sessions. These findings were equivalent for both motion tracking and video-based assessment parameters.

Table 6f Comparison of performance at the stay suture during pre- and post-course assessments, and with experienced surgeons. Data is reported as medians, with interquartile range in parentheses.

	Pre-course (n=16)	*Post-course (n=16)*	*P (pre vs. post)*	*Experienced surgeons (n=8)*	*P (pre vs. exp)*	*P (post vs. exp)*
Total time (seconds)	146 (115 – 168)	150 (118 – 190)	0.642	80 (71 – 125)	0.003	0.002
Total path length (metres)	15 (12 – 17)	15 (12 – 21)	0.278	9 (7 – 13)	0.011	0.006
Total number of movements	94 (83 – 110)	100 (77 – 134)	0.552	54 (51 – 69)	0.002	0.005
Checklist score (out of 29)	14 (12 – 19)	14 (10 – 19)	0.284	20 (14 – 24)	0.011	0.008

6.5 Conclusions

Acquisition of skill for a surgical procedure should occur in a stepwise manner, ensuring proficiency is achieved on one stage prior to progressing onto the next (chapter 2). It is also sensible to suggest that the more complex the procedure, the greater the necessity to adopt this approach. Current published material pertaining to training in LRYGBP consists of audit-based reviews of the learning curve, based upon parameters such as time taken, complication and mortality rates for the whole procedure (Schauer et al., 2003;Nguyen et al., 2003;Oliak et al., 2003;Ballesta-Lopez et al., 2005;Ballantyne et al., 2005;Shikora et al., 2005). There are two problems with this approach: first, time is not an appropriate measure for assessment of technical skill(Darzi *et al.*, 1999), and second, complication or conversion rates are not solely a function of technical expertise. Outcomes may be confounded by a number of patient factors such as co-morbidities, the experience of the operating team, and post-operative ward care (Vincent *et al.*, 2004). In order for training and assessment to occur in a standardized manner, the tasks must be standardized.

This study has defined a standardized, task-based approach to training in LRYGB in the animate skills laboratory, augmented by audio-visual material. With this mode of training, it is possible to teach the elementary skills such as intracorporeal suturing, and use of mechanical stapling devices. Bringing these core elements together enabled the development of an inanimate model for assessment of advanced laparoscopic skill. The appeal of an inanimate model is that it is inexpensive, feasible, transportable, and reproducible (Fried *et al.*, 2004a). Deficiencies of inanimate models have been their lack of realism; however, the face validity of the solution-filled cadaveric porcine model was excellent, almost restoring the bowel to its in vivo appearance (figure 6b).

The model enabled surgeons to use real laparoscopic instruments, and demonstration of construct validity confirmed that performance on the model is directly related to level of experience. Importantly, this was true not only for dexterity, but also for quality of the procedure as scored by the global and procedure-specific rating scales. It is also useful to note that the inter-rater reliability between the two assessors for these scales was high. This model is thus an objective and valid measure of technical skill in performing advanced laparoscopic procedures, and may be useful for assessment of technical skill before and after a training course.

The surgeons attending the two-day course were seen to improve in their skills to perform the jejuno-jejunostomy model, but only in terms of dexterity. Their performance on both rating scales did not improve, as was the case for completion of the initial stay suture. Though surgeons on the course made significant improvements in their dexterity to complete the jejuno-jejunostomy, they were not taught the elements of intracorporeal suturing. Without this knowledge, they shall not be able to progress onto performing a full LRYGBP procedure. This drives home the need to learn elementary skills prior to progression along the training curriculum. It is the aim to include training on laparoscopic suturing techniques as a key element of future courses.

In tune with the requirements of SAGES for training in LRYGBP to occur in a stepwise manner, this work provides the initial stages of this curriculum (2003). The first step, following knowledge of the procedure, should be familiarisation to the tools used in advanced procedures, e.g. Ligasure^{TM}, harmonic scalpel, staplers and needle holders. This follows on to training and subsequent demonstration of proficiency in intracorporeal suturing techniques on simulated models, for example in a video-box trainer. The next stage is to practice on the cadaveric porcine LRYGB model, prior to progression onto anesthetized porcine models. Each stage of this curricular approach necessitates demonstration of proficiency utilizing objective assessment criteria, as has been shown for the cadaveric porcine model (Aggarwal *et al.*, 2004b).

Within this study, the impact of two-day laparoscopic training courses for LRYGBP has been objectively proven. However, it is important to appreciate that whilst the surgeons improved their levels of skill, their dexterity at the end of the course remained inferior to that of surgeons experienced in advanced laparoscopic techniques. The acquisition of skill to achieve proficiency on the model shall entail a longer period of training, prior to progression onto the full procedure. This data also enables one to audit the quality of courses in terms of faculty and resources, and also to assess the impact of modifications to future courses.

The ultimate aim of training in surgery is to endure the bulk of the learning curve on simulated tissue, in an analogous manner to the airline industry. The transfer-effectiveness ratio (TER) is a measure of how much time (and money) pilots can save in the air by training on high-fidelity computer-based simulators on the ground (Roscoe, 1971;Roscoe, 1972). In the same way as has been shown for laparoscopic cholecystectomy in chapter 5, the intention is to train a group of surgeons with this curricular approach and assess the number of procedures necessary to achieve competency at performing a jejuno-jejunostomy on patients. Comparison with the traditional manner of training can provide one with a transfer-effectiveness ratio for the use of simulation to reduce the learning curve for LRYGBP. This not only has benefits for patients, but is also more time and cost-effective than acquiring skills on patients (Bridges and Diamond, 1999). Nonetheless, there shall always be a learning period within the operating room, and it is not the intention to disregard this. The aim is simply to shorten this period with prior training outside the operating room.

A limitation of this study was the need to ensure all surgeons performed the task in a standardised manner. This was necessary to ensure motion analysis parameters were valid, though may have skewed the results of surgeons unfamiliar with a particular instrument or technique. In order to reduce this as a limiting factor, all surgeons were provided with time to familiarize with the task and instruments. In terms of inexperienced operators, it is believed that initial training must define a safe approach in one technique, after which other techniques can be incorporated, e.g. suturing with assisted devices such as Endostitch^{TM} (U.S. Surgical, Norwalk, CT, USA). It may also be suggested that participants on the course simply improved their performance on the model. If this were true, then one would also have seen improvements in suturing skills.

The proposition is application of this curriculum for commencement of training in advanced laparoscopic procedures, particularly for LRYGBP. The paradigm is simple, cheap, flexible and portable, utilizing tools which are available in the majority of training centres. Furthermore, a new training paradigm of realistic simulation on cadaveric porcine models has been developed, which may be applicable to other advanced laparoscopic procedures.

The use of motion analysis may suggest a limitation to its widespread use, though video analysis scores could be used alone, enabling a global delivery of this training program. As an adjunct, it may be permissible to add end-product analysis to the assessment process, e.g. burst-testing of the anastomosis. This was not performed, though is under consideration for future studies.

Though not mentioned in this chapter until this point, it is also essential to consider the skills required beyond the operative environment for any aspiring bariatric surgeon, i.e. knowledge-based learning (chapter 2.3.1). Surgeons must be well versed in the pre- and post-operative management of such patients, and in the development of a team-based approach to setting up a bariatric surgical program. It is envisaged that the proposals from the American College of Surgeons and American Society of Bariatric Surgery to credential sites to perform and train individuals in bariatric surgical techniques shall enable the delivery of this aim (Pratt *et al.*, 2006).

Chapter 7

CONCLUSION

7.1 Aims of Thesis

The advent of laparoscopic surgical procedures in the late 1980s has led to improved patient outcomes in terms of reduced pain, shorter hospital stay and smaller scars. However, the introduction of these techniques required surgeons to learn new skills, unfortunately leading to a higher incidence of complications at commencement of the learning curve. It was soon evident that structured training programs were required, with in-built and objective measures of competence. Nonetheless, a proficiency-based laparoscopic surgical training curriculum does not exist, nor a methodology for how to develop such a training program.

The aims of this thesis were to:

1. Develop a standardized and universal methodology for acquisition of technical surgical skills.
2. Define a proficiency-based curriculum on a virtual reality simulator for laparoscopic skills acquisition.
3. Investigate the feasibility, validity and reliability of laparoscopic skills assessment in the operating theatre.
4. Prove the effectiveness of a virtual reality training curriculum in improving laparoscopic skill on real procedures.
5. Construct an evidence-based training program for an advanced laparoscopic surgical procedure.

7.2 The Development of a Proficiency-Based Curriculum for Laparoscopic Surgery

Surgical training is an experiential process, with skills acquired through observation, practice and feedback, preferably in a graded manner. The past decade has examined the modes of learning in surgical specialties, fuelled by the demand for experienced individuals to practice new technologies such as laparoscopy and endoscopy. This has not occurred from a theoretical base, rather through the development and application of new modes of training outside the operating theatre, i.e. simulation-based skills acquisition.

The aim of this thesis was to develop a methodology, or framework, which is simplistic, feasible and generic to any branch of medicine for acquisition of technical skill. The framework was also aligned with respected theories of motor skill learning, many of which were published in the 1970s.

The intention is to provide a step-by-step process through which medical educators can define their own training program(s). Though prescriptive, the framework is not intended to be exclusive, but one mode of achieving this aim. The current paucity of evidence-based data on the development of curricula for medical practitioners leads to a lack of standardization between training programs locally, nationally and internationally. This depth of variability subsequently transfers to differential definitions of competence for independent practice, with the possibility of variable surgical outcomes.

There are no scientific studies that describe the 'best' way to train surgical providers, how much training they need, and how to keep them technically proficient. The systematic training and assessment of technical skills (STATS) framework seeks to provide a structured technical skills training curriculum. This is based upon initial laboratory-based training, with transfer of skills to the operative environment, with in-built performance measures. Furthermore, a task-based approach is employed which can ensure not only a graded and step-wise approach to skills acquisition, but also an opportunity to identify, define and correct errors.

The notion of key tasks and how to integrate them into the development of technical skills training curriculum was demonstrated through the use of two virtual reality simulation devices for laparoscopic skills acquisition. Numerous studies had presented the construct validity and training potential of both tools, though the mode of acquiring these skills had not been presented, i.e. which task, how often, for how long, etc. The companies producing the tools had set some proficiency levels on the tasks (i.e. pass/fail), but there was no evidence behind the definition of a pass or fail. This was unsatisfactory.

In accordance with the methodological framework presented in this thesis, training was performed on all tasks, and valid assessment parameters defined. The investigation into learning curves for each task enabled the definition of the 'key tasks' which could be used for testing of proficiency. Once again, proficiency variables depended upon those that were deemed to be construct valid, and benchmark values defined upon the scores of experienced surgeons. A graded approach was also employed, moving from easy to medium to hard levels, initially defined through the study with MIST-VR in terms of whether training on all tasks, or just the two 'key tasks' was superior. Thus, a valid, evidence-based and step-wise training curriculum could be prescribed. In a similar manner, it should be possible to define such a curriculum on any other simulator device for any other surgical skill, i.e. endoscopy, bronchoscopy, endovascular procedures, arthroscopy, etc.

The added benefit of using virtual reality simulators is their ability to measure performance instantly and objectively. Furthermore, all tasks are standardized, removing the influence of patient and disease variability. The proposal is to use such curricula for the initial part of a laparoscopic training program. The LapSim training curriculum also incorporated a procedure-specific module, that of Calot's triangle dissection. Indeed, a similar process has been undertaken for incorporation of laparoscopic salpingectomy to the training curriculum for laparoscopic gynaecological surgeons (Aggarwal et al., 2006c). This can be slotted into place instead of Calot's triangle dissection. Further work seeks to extend the process to higher level tasks on virtual reality simulators as they are developed.

The important aspect of these curricula is the development of a proficiency-based paradigm to skills acquisition. The majority of studies to date have developed a training program based upon time spent in the skills laboratory, or number of sessions attended. Human performance is variable, and in order to ensure that all trainees achieve similar levels of skill, some will need to train for longer periods than others. Though self-evident, the amount of work required to develop a proficiency-based training program is far greater than that for a time-based system, in which skills acquisition is assumed rather than confirmed.

The next step within the STATS framework is to ensure transfer of skills from the laboratory to the operative environment. In order for this to occur, objective measures of performance must be in place not only during training, but also for assessment in the clinical domain. This was thus the aim of chapter 4, whereby a video-based motion tracking system (ROVIMAS) was scrutinized for assessment of laparoscopic surgical skill within the operating theatre. The initials problems of patient and disease variability were reduced through careful (and standardized) patient selection. This is important; otherwise the device is merely measuring the difficulty of the operative procedure rather than the technical skills of the operating surgeon.

In terms of feasibility of operative skills assessment, the ultimate quest is for a device which can objectively, instantly, reliably and with demonstrable validity provide evidence of one's operative performance. This is in analogy with the virtual reality simulators, which have been shown to perform this task on standardized simulation-based models. The aim was to get as close to this mode as possible within the real environment.

The integration of motion and video-based information provides instant information with regard to dexterity, together with an assessment of the quality of the procedure performed. Once again, in alignment to the STATS framework, pre-defined tasks and the whole procedure were assessed separately. This was useful as time taken was the only motion tracking variable which demonstrated construct validity, whereas Calot's triangle dissection proved significant differences between experienced and inexperienced surgeons for time taken, total path length and number of movements. This may signify that Calot's triangle dissection is the most challenging part of the procedure, and indeed a suitable assessment task. Certainly, on the basis of this study, it is the most valid and reliable for skills assessment. Indeed there were also useful correlations between the scores from video-based assessment and dexterity parameters.

The inference from this study was that whilst the video-based motion tracking device is feasible, valid and reliable for laparoscopic skills assessment in the operating theatre, perhaps a hierarchy of skills assessment devices could be developed. A crude measure is dexterity, and if individuals fail to achieve a level of proficiency, then video-based assessment is appropriate. The mode in which to identify this is to calculate the sensitivity and specificity of dexterity versus video-based assessment, following definition of the benchmark levels of skill on a real laparoscopic cholecystectomy. This needs to be tempered by the time and resources required for video-based skills assessment.

The overall aim of this chapter was to enable the use of a skills assessment system which could then be used to test the ability of the previously defined virtual reality training curriculum to lead to improved technical performance on real cases.

In the subsequent study, a comparison was made between traditional and simulation-based training strategies, with technical performance as the outcome of interest. In order to maintain standardization of procedural performance, the 'real' procedure was a porcine laparoscopic cholecystectomy. Virtual reality trained subjects had a shorter and flatter learning curve than control subjects. Furthermore, the VR-trained group managed to achieve skills equivalent to experienced surgeons for three out of four of the assessment parameters, a feat not matched by the control group. Simulation-based training was also more efficient, with every minute spent on the simulator equivalent to 2.28 minutes on the real procedure. This is not only time, but cost-efficient.

Though extremely encouraging, and the first study of its kind, an important criticism of virtual reality simulators is answered. There is concern that simulation-based training will become a replacement for the traditional mode of learning in the operating theatre, leading to the production of technically proficient, though situationally unaware surgeons. This study shows that regardless of simulation-based training to expert proficiency levels, there is a learning curve on the real procedure. The role of laboratory-based training is not to produce experienced surgeons, but rather to aid surgeons in their quest along that route. Simulation-based training is an adjunct to, and not a replacement for, the traditional mode of acquiring skills in the operating theatre. It enables trainees to adopt a shorter and flatter learning curve on real cases, leading not only to more efficient training programs, but also ones that are proficiency-based and with an education bias.

A further point of significant importance is that it was not the simulator which led to this improvement in performance. In fact, it is the curriculum which applies the simulator as a training tool which is effective. The curriculum was stepwise, task-based and graduated. Trainees did not proceed onto the next stage until they had performed to a satisfactory level at the current stage of training. It may be said that training is thus personalized to the individual, rather than upon the group as a whole.

The final chapter was an attempt to draw together previous work within the thesis for the development of a proficiency-based training curriculum for an advanced laparoscopic surgical procedure. There is great interest in the development of laparoscopic bariatric surgery, and there is a real need for a training program to ensure patients are provided with technically proficient surgeons. The STATS framework was applied, as described.

The defined procedure of laparoscopic Roux-en-Y gastric bypass (LRYGBP) was divided into four tasks. The key task, laparoscopic jejuno-jejunostomy, incorporated a number of skills common to the entire procedure. The challenge was then to produce a standardized, feasible and valid model for skills assessment. Through trial and error, a cadaveric porcine model was proven to be effective for skills assessment. The ROVIMAS assessment device was used for this purpose, together with the modified OSATS global rating scale. As an adjunct, a procedure-specific rating scale was developed. Not only did this confirm the construct validity of the model, but it may also be a useful tool for formative assessment of performance.

Though the STATS framework was implemented to test the transfer of skill from the laboratory to the operative environment, this study went on to define a further use of an assessment model. The acquisition of skill during a two-day training course was objectively assessed through pre- and post-testing of participants on the model. This was not only useful at an individual level in terms of skills performance, but also enables audit of individual courses. For example, an assessment of laparoscopic suturing skills revealed no improvements between pre- and post-course testing. This is a skill which must be mastered in order to perform LRYGBP procedures. Testing in this manner can thus inform course organizers of either the need to teach this skill during subsequent courses, or set entry criteria to the course which include demonstration of proficiency at laparoscopic suturing. This returns not only to the task-based model of training, whereby skills acquisition must occur in a stepwise manner, but also reinforces Ericsson's model of deliberate practice. Though all individuals as a group improved their performance at the laparoscopic jejuno-jejunostomy, deliberate practice at laparoscopic suturing shall lead to further improvements in the former task.

7.3 Concluding Remarks

The development and application of a framework for skills training and assessment has been described. Proficiency-based training, in a stepwise manner is the key to the definition of an effective curriculum. The acquisition of basic and advanced laparoscopic skills, together with objective and valid testing of technical outcomes has been performed.

Such tools as described can address the increasingly limited opportunities for technical training and assessment that are offered to doctors, not only during training but throughout their careers. It is no longer necessary to educate students, residents, and practicing physicians in a system that relies on chance opportunities for learning new skills. Simulation allows for risk-free training in technical skills. For the first time, a proficiency-based curriculum can make the actual level of skill rather than a predetermined period of time the primary factor in physicians' progression up the training ladder, ensuring that patients are cared for by doctors with expertise in the procedures they perform.

Although simulations alone cannot improve the quality of health care, they do significantly advance clinical education, especially when combined with enriched curricular and educational environments such as virtual operating suites, and lead to enhanced clinical reasoning and professionalism.

REFERENCES

References Cited

2003, Guidelines for institutions granting bariatric privileges utilizing laparoscopic techniques. Society of American Gastrointestinal Endoscopic Surgeons (SAGES) and the SAGES Bariatric Task Force: Surg Endosc., v. 17, no. 12, p. 2037-2040.

www.misfellowshipcouncil.org. 2005.
Ref Type: Internet Communication

2006, STEP™ Surgeons in Training Education Programme, The Royal College of Surgeons of England,
<http://www.rcseng.ac.uk/stepfdn_old/index.html/view?searchterm=step>,

Quarterly council, 1946. Minutes of Council 1945-1947. 447-461. 1947. London, Royal College of Surgeons of England.
Ref Type: Report

1998, Integrating advanced laparoscopy into surgical residency training. Society of American Gastrointestinal Endoscopic Surgeons (SAGES): Surg Endosc., v. 12, no. 4, p. 374-376.

Regulations relating to the examinations for the diploma of fellow. Minutes of Council 1941-1943. 413-423. 1943. London, Royal College of Surgeons of England.
Ref Type: Report

1991, A prospective analysis of 1518 laparoscopic cholecystectomies. The Southern Surgeons Club: N.Engl.J.Med., v. 324, no. 16, p. 1073-1078.

Aggarwal, R., and A. Darzi, 2006, Technical-skills training in the 21st century: N.Engl.J.Med., v. 355, no. 25, p. 2695-2696.

Aggarwal, R., T. P. Grantcharov, J. R. Eriksen, D. Blirup, V. B. Kristiansen, P. Funch-Jensen, and A. Darzi, 2006a, An evidence-based virtual reality training program for novice laparoscopic surgeons: Ann.Surg., v. 244, no. 2, p. 310-314.

Aggarwal, R., J. Hance, and A. Darzi, 2004a, Surgical education and training in the new millennium: Surg.Endosc., v. 18, no. 10, p. 1409-1410.

Aggarwal, R., J. Hance, S. Undre, J. Ratnasothy, K. Moorthy, A. Chang, and A. Darzi, 2006b, Training junior operative residents in laparoscopic suturing skills is feasible and efficacious: Surgery, v. 139, no. 6, p. 729-734.

Aggarwal, R., K. Moorthy, and A. Darzi, 2004b, Laparoscopic skills training and assessment: Br.J Surg, v. 91, no. 12, p. 1549-1558.

Aggarwal, R., A. Tully, T. Grantcharov, C. R. Larsen, T. Miskry, A. Farthing, and A. Darzi, 2006c, Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy: BJOG., v. 113, no. 12, p. 1382-1387.

Aggarwal, R., S. Undre, K. Moorthy, C. Vincent, and A. Darzi, 2004c, The simulated operating theatre: comprehensive training for surgical teams: Qual.Saf Health Care, v. 13 Suppl 1, p. i27-i32.

Ali, M. R., Y. Mowery, B. Kaplan, and E. J. DeMaria, 2002, Training the novice in laparoscopy. More challenge is better: Surg.Endosc., v. 16, no. 12, p. 1732-1736.

Annett, J., D. Cunningham, and P. Mathias-Jones, 2000, A method for measuring team skills: Ergonomics, v. 43, no. 8, p. 1076-1094.

Aukstakalnis, S., and D. Blatner, 1992,Aukstakalnis, S., and D. Blatner Silicon mirage: the art and science of virtual reality: Berkeley, California, Peachpit Press.

Ausubel, D. P., J. S. Novak, and H. Hanesian, 1978,Ausubel, D. P., J. S. Novak, and H. Hanesian Educational psychology: a cognitive view: New York, Holt, Rinehart and Winston.

Babineau, T. J., J. Becker, G. Gibbons, S. Sentovich, D. Hess, S. Robertson, and M. Stone, 2004, The "cost" of operative training for surgical residents: Arch.Surg., v. 139, no. 4, p. 366-369.

Bailey, R. W., A. L. Imbembo, and K. A. Zucker, 1991, Establishment of a laparoscopic cholecystectomy training program: Am.Surg., v. 57, no. 4, p. 231-236.

Ballantyne, G. H., D. Ewing, R. F. Capella, J. F. Capella, D. Davis, H. J. Schmidt, A. Wasielewski, and R. J. Davies, 2005, The learning curve measured by operating times for laparoscopic and open gastric bypass: roles of surgeon's experience, institutional experience, body mass index and fellowship training: Obes.Surg, v. 15, no. 2, p. 172-182.

Ballesta-Lopez, C., I. Poves, M. Cabrera, J. A. Almeida, and G. Macias, 2005, Learning curve for laparoscopic Roux-en-Y gastric bypass with totally hand-sewn anastomosis: analysis of first 600 consecutive patients: Surg Endosc., v. 19, no. 4, p. 519-524.

Bann, S., I. M. Davis, K. Moorthy, Y. Munz, J. Hernandez, M. Khan, V. Datta, and A. Darzi, 2005, The reliability of multiple objective measures of surgery and the role of human performance: Am J.Surg, v. 189, no. 6, p. 747-752.

Bannister, S. L., R. I. Hilliard, G. Regehr, and L. Lingard, 2003, Technical skills in paediatrics: a qualitative study of acquisition, attitudes and assumptions in the neonatal intensive care unit: Med.Educ., v. 37, no. 12, p. 1082-1090.

Barnes, R. W., N. P. Lang, and M. F. Whiteside, 1989, Halstedian technique revisited. Innovations in teaching surgical skills: Ann.Surg., v. 210, no. 1, p. 118-121.

Beard, J. D., B. C. Jolly, D. I. Newble, W. E. Thomas, J. Donnelly, and L. J. Southgate, 2005, Assessing the technical skills of surgical trainees: Br.J Surg., v. 92, no. 6, p. 778-782.

Berci,G. Editorial comment. Am.J.Surg. 160, 488-489. 1990.
Ref Type: Journal (Full)

Berci, G., and J. M. Sackier, 1991, The Los Angeles experience with laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 382-384.

Bevan, P. G., 1986, Craft workshops in surgery: Br.J.Surg., v. 73, no. 1, p. 1-2.

Bevan, P. G., 1997, The first anastomosis workshop, March 1981: Ann.R.Coll.Surg.Engl., v. 79, no. 4 Suppl, p. 168-169.

Blum, M. G., T. W. Powers, and S. Sundaresan, 2004, Bronchoscopy simulator effectively prepares junior residents to competently perform basic clinical bronchoscopy: Ann.Thorac.Surg, v. 78, no. 1, p. 287-291.

Bridges, M., and D. L. Diamond, 1999, The financial impact of teaching surgical residents in the operating room: Am.J Surg, v. 177, no. 1, p. 28-32.

Broe, D., P. F. Ridgway, S. Johnson, S. Tierney, and K. C. Conlon, 2006, Construct validation of a novel hybrid surgical simulator: Surg.Endosc., v. 20, no. 6, p. 900-904.

Brunner, W. C., J. R. Korndorffer, Jr., R. Sierra, N. N. Massarweh, J. B. Dunne, C. L. Yau, and D. J. Scott, 2004, Laparoscopic virtual reality training: are 30 repetitions enough?: J.Surg.Res., v. 122, no. 2, p. 150-156.

Bulstrode, C., and G. Holsgrove, 1996, Education for educating surgeons: BMJ, v. 312, no. 7027, p. 326-327.

Byrne, P., 1994, Teaching laparoscopic surgery. Practice on live animals is illegal: BMJ, v. 308, no. 6941, p. 1435.

Cao, C. G., C. L. MacKenzie, J. A. Ibbotson, L. J. Turner, N. P. Blair, and A. G. Nagy, 1999, Hierarchical decomposition of laparoscopic procedures: Stud.Health Technol.Inform., v. 62, p. 83-89.

Carter, F. J., M. P. Schijven, R. Aggarwal, T. Grantcharov, N. K. Francis, G. B. Hanna, and J. J. Jakimowicz, 2005, Consensus guidelines for validation of virtual reality surgical simulators: Surg.Endosc., v. 19, no. 12, p. 1523-1532.

Champion, H. R., and A. G. Gallagher, 2003, Surgical simulation - a 'good idea whose time has come': Br.J.Surg., v. 90, no. 7, p. 767-768.

Chaudhry, A., C. Sutton, J. Wood, R. Stone, and R. McCloy, 1999, Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface: Ann.R.Coll.Surg.Engl., v. 81, no. 4, p. 281-286.

Clayden, G. S., 1994, Teaching laparoscopic surgery. Preliminary training on animals is essential: BMJ, v. 309, no. 6950, p. 342.

Cohen, R., R. K. Reznick, B. R. Taylor, J. Provan, and A. Rothman, 1990, Reliability and validity of the objective structured clinical examination in assessing surgical residents: Am.J.Surg., v. 160, no. 3, p. 302-305.

Coleman, R. L., and C. Y. Muller, 2002, Effects of a laboratory-based skills curriculum on laparoscopic proficiency: a randomized trial: Am.J.Obstet.Gynecol., v. 186, no. 4, p. 836-842.

Colquitt, J., A. Clegg, M. Sidhu, and P. Royle, 2003, Surgery for morbid obesity: Cochrane.Database.Syst.Rev., no. 2, p. CD003641.

Crothers, I. R., A. G. Gallagher, N. McClure, D. T. James, and J. McGuigan, 1999, Experienced laparoscopic surgeons are automated to the "fulcrum effect": an ergonomic demonstration: Endoscopy, v. 31, no. 5, p. 365-369.

Cuschieri, A., 1992, The dust has settled--let's sweep it clean: training in minimal access surgery: J.R.Coll.Surg.Edinb., v. 37, no. 4, p. 213-214.

Cuschieri, A., 2003, Medical error, incidents, accidents and violations: Min Inv Ther Allied Technol, v. 12, no. 3-4, p. 111-120.

Cuschieri, A., 1995, Whither minimal access surgery: tribulations and expectations: Am.J.Surg., v. 169, no. 1, p. 9-19.

Cuschieri, A., G. Berci, and C. K. McSherry, 1990, Laparoscopic cholecystectomy: Am.J.Surg., v. 159, no. 3, p. 273.

Cuschieri, A., F. Dubois, J. Mouiel, P. Mouret, H. Becker, G. Buess, M. Trede, and H. Troidl, 1991, The European experience with laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 385-387.

Cuschieri, A., N. Francis, J. Crosby, and G. B. Hanna, 2001, What do master surgeons think of surgical competence and revalidation?: Am.J.Surg., v. 182, no. 2, p. 110-116.

Cuschieri, A., F. A. Gleeson, R. M. Harden, and R. A. Wood, 1979, A new approach to a final examination in surgery. Use of the objective structured clinical examination: Ann.R.Coll.Surg.Engl., v. 61, no. 5, p. 400-405.

Darzi, A., 1996,Darzi, A. Laparoscopic Cholecystectomy Course Handbook: London, Royal College of Surgeons of England.

Darzi, A., V. Datta, and S. Mackay, 2001, The challenge of objective assessment of surgical skill: Am.J.Surg., v. 181, no. 6, p. 484-486.

Darzi, A., S. Smith, and N. Taffinder, 1999, Assessing operative skill. Needs to become more objective: BMJ, v. 318, no. 7188, p. 887-888.

Dath, D., G. Regehr, D. Birch, C. Schlachta, E. Poulin, J. Mamazza, R. Reznick, and H. M. MacRae, 2004, Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills: Surg.Endosc., v. 18, no. 12, p. 1800-1804.

Datta, V., S. Mackay, M. Mandalia, and A. Darzi, 2001, The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model: J.Am.Coll.Surg., v. 193, no. 5, p. 479-485.

Debas, H. T. et al., 2005, American Surgical Association Blue Ribbon Committee Report on Surgical Education: 2004: Ann.Surg., v. 241, no. 1, p. 1-8.

Dempere-Marco, L., X. P. Hu, S. M. Ellis, D. M. Hansell, and G. Z. Yang, 2006, Analysis of visual search patterns with EMD metric in normalized anatomical space: IEEE Trans.Med.Imaging, v. 25, no. 8, p. 1011-1021.

Dent, T. L., 1991, Training, credentialling, and granting of clinical privileges for laparoscopic general surgery: Am.J.Surg., v. 161, no. 3, p. 399-403.

Dent, T. L., Ponsky JL, and G. Berci, 1991, Minimal access general surgery: the dawn of a new era: Am.J.Surg., v. 161, p. 323.

Derossis, A. M., G. M. Fried, M. Abrahamowicz, H. H. Sigman, J. S. Barkun, and J. L. Meakins, 1998, Development of a model for training and evaluation of laparoscopic skills: Am.J.Surg., v. 175, no. 6, p. 482-487.

Dosis, A., R. Aggarwal, F. Bello, K. Moorthy, Y. Munz, D. Gillies, and A. Darzi, 2005, Synchronized video and motion analysis for the assessment of procedures in the operating theater: Arch.Surg, v. 140, no. 3, p. 293-299.

Dubois, F., G. Berthelot, and H. Levard, 1989, [Cholecystectomy by coelioscopy]: Presse Med., v. 18, no. 19, p. 980-982.

Dubois, F., P. Icard, G. Berthelot, and H. Levard, 1990, Coelioscopic cholecystectomy. Preliminary report of 36 cases: Ann.Surg., v. 211, no. 1, p. 60-62.

Duffy, A. J., N. J. Hogle, H. McCarthy, J. I. Lew, A. Egan, P. Christos, and D. L. Fowler, 2005, Construct validity for the LAPSIM laparoscopic surgical simulator: Surg.Endosc., v. 19, no. 3, p. 401-405.

Dunn, D., R. Nair, S. Fowler, and R. McCloy, 1994, Laparoscopic cholecystectomy in England and Wales: results of an audit by the Royal College of Surgeons of England: Ann.R.Coll.Surg.Engl., v. 76, no. 4, p. 269-275.

Duthie, G. S., P. J. Drew, M. A. Hughes, R. Farouk, R. Hodson, K. R. Wedgwood, and J. R. Monson, 1998, A UK training programme for nurse practitioner flexible sigmoidoscopy and a prospective evaluation of the practice of the first UK trained nurse flexible sigmoidoscopist: Gut, v. 43, no. 5, p. 711-714.

Elliot, D. L., and D. H. Hickam, 1987, Evaluation of physical examination skills. Reliability of faculty observers and patient instructors: JAMA, v. 258, no. 23, p. 3405-3408.

Elwyn, G., and J. M. Corrigan, 2005, The patient safety story: BMJ, v. 331, no. 7512, p. 302-304.

Emam, T. A., G. Hanna, and A. Cuschieri, 2002, Ergonomic principles of task alignment, visual display, and direction of execution of laparoscopic bowel suturing: Surg.Endosc., v. 16, no. 2, p. 267-271.

Ericsson, K. A., 2004, Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains: Acad.Med., v. 79, no. 10 Suppl, p. S70-S81.

Eriksen, J. R., and T. Grantcharov, 2005, Objective assessment of laparoscopic skills using a virtual reality stimulator: Surg.Endosc., v. 19, no. 9, p. 1216-1219.

Eubanks, T. R., R. H. Clements, D. Pohl, N. Williams, D. C. Schaad, S. Horgan, and C. Pellegrini, 1999, An objective scoring system for laparoscopic cholecystectomy: J.Am.Coll.Surg., v. 189, no. 6, p. 566-574.

European Association of Endoscopic Surgeons, 1994, Training and assessment of competence: Surg.Endosc., v. 8, no. 6, p. 721-722.

Figert, P. L., A. E. Park, D. B. Witzke, and R. W. Schwartz, 2001, Transfer of training in acquiring laparoscopic skills: J.Am.Coll.Surg., v. 193, no. 5, p. 533-537.

Fitts AM, and Posner MI, 1967,Fitts AM, and Posner MI Human Performance: Belmont, CA, Brooks-Cole.

Flowers, J. L., R. W. Bailey, W. A. Scovill, and K. A. Zucker, 1991, The Baltimore experience with laparoscopic management of acute cholecystitis: Am.J.Surg., v. 161, no. 3, p. 388-392.

Francis, N. K., G. B. Hanna, and A. Cuschieri, 2001, Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks: Arch.Surg., v. 136, no. 1, p. 40-43.

Francis, N. K., G. B. Hanna, and A. Cuschieri, 2002, The performance of master surgeons on the Advanced Dundee Endoscopic Psychomotor Tester: contrast validity study: Arch.Surg., v. 137, no. 7, p. 841-844.

Fried, G. M., A. M. Derossis, J. Bothwell, and H. H. Sigman, 1999, Comparison of laparoscopic performance in vivo with performance measured in a laparoscopic simulator: Surg.Endosc., v. 13, no. 11, p. 1077-1081.

Fried, G. M., L. S. Feldman, M. C. Vassiliou, S. A. Fraser, D. Stanbridge, G. Ghitulescu, and C. G. Andrew, 2004a, Proving the value of simulation in laparoscopic surgery: Ann.Surg., v. 240, no. 3, p. 518-525.

Fried, M. P. et al., 2004b, Identifying and reducing errors with surgical simulation: Qual.Saf Health Care, v. 13 Suppl 1, p. i19-i26.

Gadacz, T. R., and M. A. Talamini, 1991, Traditional versus laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 336-338.

Gallagher, A. G., and C. U. Cates, 2004, Virtual reality training for the operating room and cardiac catheterisation laboratory: Lancet, v. 364, no. 9444, p. 1538-1540.

Gallagher, A. G., A. B. Lederman, K. McGlade, R. M. Satava, and C. D. Smith, 2004, Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance: Surg.Endosc., v. 18, no. 4, p. 660-665.

Gallagher, A. G., N. McClure, J. McGuigan, K. Ritchie, and N. P. Sheehy, 1998, An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills: Endoscopy, v. 30, no. 7, p. 617-620.

Gallagher, A. G., K. Richie, N. McClure, and J. McGuigan, 2001, Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality: World J.Surg., v. 25, no. 11, p. 1478-1483.

Gallagher, A. G., E. M. Ritter, and R. M. Satava, 2003a, Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training: Surg.Endosc., v. 17, no. 10, p. 1525-1529.

Gallagher, A. G., and R. M. Satava, 2002, Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures: Surg.Endosc., v. 16, no. 12, p. 1746-1752.

Gallagher, A. G., C. D. Smith, S. P. Bowers, N. E. Seymour, A. Pearson, S. McNatt, D. Hananel, and R. M. Satava, 2003b, Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures: J.Am.Coll.Surg., v. 197, no. 3, p. 479-488.

Gor, M., R. McCloy, R. Stone, and A. Smith, 2003, Virtual reality laparoscopic simulator for assessment in gynaecology: BJOG., v. 110, no. 2, p. 181-187.

Grantcharov, T. P., L. Bardram, P. Funch-Jensen, and J. Rosenberg, 2003, Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills: Am.J.Surg., v. 185, no. 2, p. 146-149.

Grantcharov, T. P., V. B. Kristiansen, J. Bendix, L. Bardram, J. Rosenberg, and P. Funch-Jensen, 2004, Randomized clinical trial of virtual reality simulation for laparoscopic skills training: Br.J.Surg., v. 91, no. 2, p. 146-150.

Grantcharov, T. P., J. Rosenberg, E. Pahle, and P. Funch-Jensen, 2001, Virtual reality computer simulation: Surg.Endosc., v. 15, no. 3, p. 242-244.

Greenhalgh, R. M., H. H. Eastcott, A. O. Mansfield, and D. E. Taylor, 1987, Aneurysm jig for anastomosis technique: Ann.R.Coll.Surg.Engl., v. 69, no. 5, p. 199-200.

Grober, E. D., S. J. Hamstra, K. R. Wanzel, R. K. Reznick, E. D. Matsumoto, R. S. Sidhu, and K. A. Jarvi, 2004a, The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures: Ann.Surg., v. 240, no. 2, p. 374-381.

Grober, E. D., S. J. Hamstra, K. R. Wanzel, R. K. Reznick, E. D. Matsumoto, R. S. Sidhu, and K. A. Jarvi, 2004b, The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures: Ann.Surg., v. 240, no. 2, p. 374-381.

Guerlain, S., R. B. Adams, F. B. Turrentine, T. Shin, H. Guo, S. R. Collins, and J. F. Calland, 2005, Assessing team performance in the operating room: development and use of a "black-box" recorder and other tools for the intraoperative environment: J.Am Coll Surg, v. 200, no. 1, p. 29-37.

Hall, J. C., C. Ellis, and J. Hamdorf, 2003, Surgeons and cognitive processes: Br.J.Surg., v. 90, no. 1, p. 10-16.

Hall, J. C., J. M. Watts, P. E. O'Brien, R. E. Dunstan, J. F. Walsh, A. H. Slavotinek, and R. G. Elmslie, 1990, Gastric surgery for morbid obesity. The Adelaide Study: Ann.Surg, v. 211, no. 4, p. 419-427.

Halsted, W. S., 1904, The training of the surgeon: Bull Johns Hopkins Hosp, v. 15, p. 267-275.

Hamdorf, J. M., and J. C. Hall, 2000, Acquiring surgical skills: Br.J.Surg., v. 87, no. 1, p. 28-37.

Hamilton, E. C., D. J. Scott, J. B. Fleming, R. V. Rege, R. Laycock, P. C. Bergen, S. T. Tesfay, and D. B. Jones, 2002, Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills: Surg.Endosc., v. 16, no. 3, p. 406-411.

Hanna, G. B., T. Drew, P. Clinch, B. Hunter, and A. Cuschieri, 1998a, Computer-controlled endoscopic performance assessment system: Surg.Endosc., v. 12, no. 7, p. 997-1000.

Hanna, G. B., S. M. Shimi, and A. Cuschieri, 1998b, Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy: Lancet, v. 351, no. 9098, p. 248-251.

Hassan, I., B. Gerdes, M. Koller, P. Langer, M. Rothmund, and A. Zielke, 2006, Clinical background is required for optimum performance with a VR laparoscopy simulator: Comput.Aided Surg., v. 11, no. 2, p. 103-106.

Hassan, I., H. Sitter, K. Schlosser, A. Zielke, M. Rothmund, and B. Gerdes, 2005, [A virtual reality simulator for objective assessment of surgeons' laparoscopic skill]: Chirurg, v. 76, no. 2, p. 151-156.

Higa, K. D., K. B. Boone, and T. Ho, 2000, Complications of the laparoscopic Roux-en-Y gastric bypass: 1,040 patients--what have we learned?: Obes.Surg, v. 10, no. 6, p. 509-513.

Hill, J., and E. S. Kiff, 1990, An abdominal wall jig for surgical craft workshops: Ann.R.Coll.Surg.Engl., v. 72, no. 6, p. 386-387.

Hyltander, A., E. Liljegren, P. H. Rhodin, and H. Lonroth, 2002, The transfer of basic skills learned in a laparoscopic simulator to the operating room: Surg.Endosc., v. 16, no. 9, p. 1324-1328.

Immenroth, M., T. Burger, J. Brenner, M. Nagelschmidt, H. Eberspacher, and H. Troidl, 2007, Mental training in surgical education: a randomized controlled trial: Ann.Surg..

Jackson B. Surgical Competence: Challenges of assessment in training and practice. 5. 1999. London, RCS & Smith and Nephew.
Ref Type: Report

Joice, P., G. B. Hanna, and A. Cuschieri, 1998, Errors enacted during endoscopic surgery--a human reliability analysis: Appl.Ergon., v. 29, no. 6, p. 409-414.

Jolly, B., and L. Rees, 1998,Jolly, B., and L. Rees Medical education in the millenium: Oxford, Oxford University Press.

Jones, H. J., R. Coggins, J. Lafuente, and C. L. de, 1999, Value of a surgical high-dependency unit: Br.J.Surg., v. 86, no. 12, p. 1578-1582.

Jordan, J. A., A. G. Gallagher, J. McGuigan, and N. McClure, 2001, Virtual reality training leads to faster adaptation to the novel psychomotor restrictions encountered by laparoscopic surgeons: Surg.Endosc., v. 15, no. 10, p. 1080-1084.

Katz, R., A. Hoznek, L. Salomon, P. Antiphon, T. A. de la, and C. C. Abbou, 2005, Skill assessment of urological laparoscopic surgeons: can criterion levels of surgical performance be determined using the pelvic box trainer?: Eur.Urol., v. 47, no. 4, p. 482-487.

Kirwan, W. O., T. K. Kaar, and R. Waldron, 1991, Starting laparoscopic cholecystectomy--the pig as a training model: Ir.J.Med.Sci., v. 160, no. 8, p. 243-246.

Kneebone, R., and A. Darzi, 2005, New professional roles in surgery: BMJ, v. 330, no. 7495, p. 803-804.

Kneebone, R. L., D. Nestel, K. Moorthy, P. Taylor, S. Bann, Y. Munz, and A. Darzi, 2003, Learning the skills of flexible sigmoidoscopy - the wider perspective: Med.Educ., v. 37 Suppl 1, p. 50-58.

Kneebone, R. L., W. Scott, A. Darzi, and M. Horrocks, 2004, Simulation and clinical practice: strengthening the relationship: Med.Educ., v. 38, no. 10, p. 1095-1102.

Kolb, D. A., 1984,Kolb, D. A. Experiential learning: experience as a source of learning and development: Englewood Cliffs, New Jersey, Prentice Hall.

Korndorffer, J. R., Jr., D. J. Scott, R. Sierra, W. C. Brunner, J. B. Dunne, D. P. Slakey, M. C. Townsend, and R. L. Hewitt, 2005, Developing and testing competency levels for laparoscopic skills training: Arch.Surg, v. 140, no. 1, p. 80-84.

Korndorffer, J. R., Jr., D. Stefanidis, and D. J. Scott, 2006, Laparoscopic skills laboratories: current assessment and a call for resident training standards: Am J Surg, v. 191, no. 1, p. 17-22.

Kothari, S. N., W. C. Boyd, C. A. Larson, H. L. Gustafson, P. J. Lambert, and M. A. Mathiason, 2005, Training of a minimally invasive bariatric surgeon: are laparoscopic fellowships the answer?: Obes.Surg, v. 15, no. 3, p. 323-329.

Kothari, S. N., B. J. Kaplan, E. J. DeMaria, T. J. Broderick, and R. C. Merrell, 2002, Training in laparoscopic suturing skills using a new computer-based virtual reality simulator (MIST-VR) provides results comparable to those with an established pelvic trainer system: J.Laparoendosc.Adv.Surg.Tech.A, v. 12, no. 3, p. 167-173.

Krummel, T. M., 1998, Surgical simulation and virtual reality: the coming revolution: Ann.Surg., v. 228, no. 5, p. 635-637.

Kumar, B. D., Y. Munz, K. Moorthy, and A. Darzi, 2003, How can handful of water assess basic laparoscopic skills?: Ann.R.Coll.Surg.Engl., v. 85, no. 6, p. 426-427.

Lamata, P., E. J. Gomez, F. Bello, R. L. Kneebone, R. Aggarwal, and F. Lamata, 2006, Conceptual framework for laparoscopic VR simulators: IEEE Comput.Graph.Appl., v. 26, no. 6, p. 69-79.

Langelotz, C., M. Kilian, C. Paul, and W. Schwenk, 2005, LapSim virtual reality laparoscopic simulator reflects clinical experience in German surgeons: Langenbecks Arch.Surg., v. 390, no. 6, p. 534-537.

Larson, J. L., R. G. Williams, J. Ketchum, M. L. Boehler, and G. L. Dunnington, 2005, Feasibility, reliability and validity of an operative performance rating system for evaluating surgery residents: Surgery, v. 138, no. 4, p. 640-647.

Larsson, A., 2001, An open and flexible framework for computer aided surgical training: Stud.Health Technol.Inform., v. 81, p. 263-265.

Liangand, K. Y., and S. L. Zeger, 1986, Longitudinal data analysis using generalized linear models: Biometrika, v. 73, p. 13-22.

Lin, E., S. Szomstein, T. Addasi, L. Galati-Burke, J. W. Turner, and H. I. Tiszenkel, 2003, Model for teaching laparoscopic colectomy to surgical residents: Am.J.Surg., v. 186, no. 1, p. 45-48.

Macintyre, I. M., and R. G. Wilson, 1993, Laparoscopic cholecystectomy: Br.J.Surg., v. 80, no. 5, p. 552-559.

Mackay, S., P. Morgan, V. Datta, A. Chang, and A. Darzi, 2002, Practice distribution in procedural skills training: a randomized controlled trial: Surg.Endosc., v. 16, no. 6, p. 957-961.

MacKenzie, C. L., J. A. Ibbotson, C. G. Cao, and A. J. Lomax, 2001, Hierarchical decomposition of laparoscopic surgery: a human factors approach to investigating the operating room environment: Min Invas Ther & Allied Technol, v. 10, no. 3, p. 121-127.

Macmillan, A. I., and A. Cuschieri, 1999, Assessment of innate ability and skills for endoscopic manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: predictive and concurrent validity: Am.J.Surg., v. 177, no. 3, p. 274-277.

Madan, A. K., C. T. Frantzides, and L. M. Sasso, 2005, Laparoscopic baseline ability assessment by virtual reality: J.Laparoendosc.Adv.Surg.Tech.A, v. 15, no. 1, p. 13-17.

Maggard, M. A. et al., 2005, Meta-analysis: surgical treatment of obesity: Ann.Intern.Med, v. 142, no. 7, p. 547-559.

Majeed, A. W., M. W. Reed, and A. G. Johnson, 1992, Simulated laparoscopic cholecystectomy: Ann.R.Coll.Surg.Engl., v. 74, no. 1, p. 70-71.

Martin, J. A., G. Regehr, R. Reznick, H. MacRae, J. Murnaghan, C. Hutchison, and M. Brown, 1997, Objective structured assessment of technical skill (OSATS) for surgical residents: Br.J.Surg., v. 84, no. 2, p. 273-278.

Mason, E. E., and C. Ito, 1967, Gastric bypass in obesity: Surg Clin.North Am., v. 47, no. 6, p. 1345-1351.

Mayer, R. E., 1979, Can advance organisers influence meaningful learning?: Review of Educational Research, v. 49, p. 371-383.

McDougall, E. M., F. A. Corica, J. R. Boker, L. G. Sala, G. Stoliar, J. F. Borin, F. T. Chu, and R. V. Clayman, 2006, Construct validity testing of a laparoscopic surgical simulator: J.Am.Coll.Surg., v. 202, no. 5, p. 779-787.

McGovern, K. T., and L. T. McGovern, 1994, The virtual clinic, a virtual reality surgical clinic: Virt Real Worl, v. Mar-Apr, p. 41-44.

McNatt, S. S., and C. D. Smith, 2001, A computer-based laparoscopic skills assessment device differentiates experienced from novice laparoscopic surgeons: Surg.Endosc., v. 15, no. 10, p. 1085-1089.

Mertz, H., and S. Gautam, 2004, The learning curve for EUS-guided FNA of pancreatic cancer: Gastrointest.Endosc., v. 59, no. 1, p. 33-37.

Moore, M. J., and C. L. Bennett, 1995, The learning curve for laparoscopic cholecystectomy. The Southern Surgeons Club: Am.J Surg, v. 170, no. 1, p. 55-59.

Moorthy K, Munz Y, Dosis A, Bello F, A Darzi. Motion analysis in the training and assessment of minimally invasive surgery. Min Invas Ther & Allied Technol 12[3-4], 137-142. 2003.
Ref Type: Journal (Full)

Moorthy, K., Y. Munz, S. Adams, V. Pandey, and A. Darzi, 2005, A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre: Ann.Surg, v. 242, no. 5, p. 631-639.

Moorthy, K., Y. Munz, A. Dosis, F. Bello, A. Chang, and A. Darzi, 2004, Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity: Surg Endosc., v. 18, no. 11, p. 1608-1612.

Moorthy, K., Y. Munz, D. Forrest, V. Pandey, S. Undre, C. Vincent, and A. Darzi, 2006, Surgical crisis management skills training and assessment: a simulation[corrected]-based approach to enhancing operating room performance: Ann.Surg, v. 244, no. 1, p. 139-147.

Moorthy, K., Y. Munz, S. K. Sarker, and A. Darzi, 2003, Objective assessment of technical skills in surgery: BMJ, v. 327, no. 7422, p. 1032-1037.

Morino, M., V. Festa, and C. Garrone, 1995, Survey on Torino courses. The impact of a two-day practical course on apprenticeship and diffusion of laparoscopic cholecystectomy in Italy: Surg.Endosc., v. 9, no. 1, p. 46-48.

Mouret,P. Presentation at Society of American Gastrointestinal Surgeons, Louisville, Kentucky. 1989.
Ref Type: Conference Proceeding

Mughal, M., 1992, A cheap laparoscopic surgery trainer: Ann.R.Coll.Surg.Engl., v. 74, no. 4, p. 256-257.

Munz, Y., B. D. Kumar, K. Moorthy, S. Bann, and A. Darzi, 2004, Laparoscopic virtual reality and box trainers: is one superior to the other?: Surg.Endosc., v. 18, no. 3, p. 485-494.

Neville, E., 2003, Modernising medical careers: Clin.Med, v. 3, no. 6, p. 529-531.

Nguyen, N. T., C. Goldman, C. J. Rosenquist, A. Arango, C. J. Cole, S. J. Lee, and B. M. Wolfe, 2001a, Laparoscopic versus open gastric bypass: a randomized study of outcomes, quality of life, and costs: Ann.Surg, v. 234, no. 3, p. 279-289.

Nguyen, N. T., H. S. Ho, L. S. Palmer, and B. M. Wolfe, 2000, A comparison study of laparoscopic versus open gastric bypass for morbid obesity: J Am.Coll.Surg, v. 191, no. 2, p. 149-155.

Nguyen, N. T., S. L. Lee, C. Goldman, N. Fleming, A. Arango, R. McFall, and B. M. Wolfe, 2001b, Comparison of pulmonary function and postoperative pain after laparoscopic versus open gastric bypass: a randomized trial: J Am.Coll.Surg, v. 192, no. 4, p. 469-476.

Nguyen, N. T., R. Rivers, and B. M. Wolfe, 2003, Factors associated with operative outcomes in laparoscopic gastric bypass: J Am.Coll.Surg, v. 197, no. 4, p. 548-555.

Nguyen, N. T., and B. M. Wolfe, 2002, Laparoscopic versus open gastric bypass: Semin.Laparosc.Surg, v. 9, no. 2, p. 86-93.

Olbers, T., M. Fagevik-Olsen, A. Maleckas, and H. Lonroth, 2005, Randomized clinical trial of laparoscopic Roux-en-Y gastric bypass versus laparoscopic vertical banded gastroplasty for obesity: Br.J Surg, v. 92, no. 5, p. 557-562.

Oliak, D., G. H. Ballantyne, P. Weber, A. Wasielewski, R. J. Davies, and H. J. Schmidt, 2003, Laparoscopic Roux-en-Y gastric bypass: defining the learning curve: Surg Endosc., v. 17, no. 3, p. 405-408.

Olsen, D. O., 1991, Laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 339-344.

Patkin, M., and L. Isabel, 1995, Ergonomics, engineering and surgery of endosurgical dissection: J R.Coll.Surg.Edinb., v. 40, no. 2, p. 120-132.

Payandeh, S., A. J. Lomax, J. Dill, C. L. MacKenzie, and C. G. Cao, 2002, On defining metrics for assessing laparoscopic surgical skills in a virtual training environment: Stud.Health Technol Inform., v. 85, p. 334-340.

Pearson, A. M., A. G. Gallagher, J. C. Rosser, and R. M. Satava, 2002, Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying: Surg.Endosc., v. 16, no. 1, p. 130-137.

Pellegrini, C. A., 2002, Invited commentary: the ACGME "Outcomes Project". American Council for Graduate Medical Education: Surgery, v. 131, no. 2, p. 214-215.

Perissat, J., D. Collet, and R. Belliard, 1990, Gallstones: laparoscopic treatment--cholecystectomy, cholecystostomy, and lithotripsy. Our own technique: Surg.Endosc., v. 4, no. 1, p. 1-5.

Perissat, J., and G. C. Vitale, 1991, Laparoscopic cholecystectomy: gateway to the future: Am.J.Surg., v. 161, no. 3, p. 408.

Peters, J. H., E. C. Ellison, J. T. Innes, J. L. Liss, K. E. Nichols, J. M. Lomano, S. R. Roby, M. E. Front, and L. C. Carey, 1991, Safety and efficacy of laparoscopic cholecystectomy. A prospective analysis of 100 initial patients: Ann.Surg., v. 213, no. 1, p. 3-12.

Pickersgill, T., 2001, The European working time directive for doctors in training: BMJ, v. 323, no. 7324, p. 1266.

Pimentel, K., and K. Teixeira, 1993,Pimentel, K., and K. Teixeira Virtual reality: through the new looking glass: Blueridge, Pennsylvania, Windcrest.

Ponsky, J. L., 1991, Complications of laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 393-395.

Powers, T. W., K. M. Murayama, M. Toyama, S. Murphy, E. W. Denham, III, A. M. Derossis, and R. J. Joehl, 2002, Housestaff performance is improved by participation in a laparoscopic skills curriculum: Am.J.Surg., v. 184, no. 6, p. 626-629.

Pratt, G. M., B. McLees, and W. J. Pories, 2006, The ASBS Bariatric Surgery Centers of Excellence program: a blueprint for quality improvement: Surg Obes.Relat Dis., v. 2, no. 5, p. 497-503.

Ragunath, K., L. A. Thomas, W. Y. Cheung, P. D. Duane, and D. G. Richards, 2003, Objective evaluation of ERCP procedures: a simple grading scale for evaluating technical difficulty: Postgrad.Med.J., v. 79, no. 934, p. 467-470.

Rantanen, E. M., and D. A. Talleur, 2005, Incremental transfer and cost effectiveness of ground-based flight trainers in a university aviation program.: Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, p. 764-768.

Rattner, D. W., K. N. Apelgren, and W. S. Eubanks, 2001, The need for training opportunities in advanced laparoscopic surgery: Surg Endosc., v. 15, no. 10, p. 1066-1070.

Reason, J., 1994,Reason, J. Human Error: Cambridge, Cambridge University Press.

Reddick,EJ. Editorial comment. Am.J.Surg. 160, 488. 1990.
Ref Type: Journal (Full)

Reddick, E. J., D. Olsen, A. Spaw, D. Baird, H. Asbun, M. O'Reilly, K. Fisher, and W. Saye, 1991, Safe performance of difficult laparoscopic cholecystectomies: Am.J.Surg., v. 161, no. 3, p. 377-381.

Reddick, E. J., and D. O. Olsen, 1989, Laparoscopic laser cholecystectomy. A comparison with mini-lap cholecystectomy: Surg.Endosc., v. 3, no. 3, p. 131-133.

Regehr, G., H. MacRae, R. K. Reznick, and D. Szalay, 1998, Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination: Acad.Med., v. 73, no. 9, p. 993-997.

Reinhardt-Rutland, A. H., and A. G. Gallagher, 1996, Visual depth perception in minimally invasive surgery, in SA Robertson ed., Contemporary Ergonomics: London, Taylor Francis, p. 531-536.

Reznick, R. K., 1993, Teaching and testing technical skills: Am.J Surg., v. 165, no. 3, p. 358-361.

Reznick, R. K., and H. MacRae, 2006, Teaching surgical skills--changes in the wind: N.Engl.J.Med., v. 355, no. 25, p. 2664-2669.

Ribble, J. G., G. L. Burkett, and G. H. Escovitz, 1981, Priorities and practices of continuing medical education program directors: JAMA, v. 245, no. 2, p. 160-163.

Ro, C. Y., I. K. Toumpoulis, R. C. Ashton, Jr., T. Jebara, C. Schulman, G. J. Todd, J. J. Derose, Jr., and J. J. McGinty, 2005, The LapSim: a learning environment for both experts and novices: Stud.Health Technol.Inform., v. 111, p. 414-417.

Rogers, D. A., G. Regehr, and J. MacDonald, 2002, A role for error training in surgical technical skill instruction and evaluation: Am J Surg, v. 183, no. 3, p. 242-245.

Roscoe, S. N., 1972, A little more on incremental transfer effectiveness: Human Factors, v. 14, no. 4, p. 363-364.

Roscoe, S. N., 1971, Incremental transfer effectiveness: Human Factors, v. 13, no. 6, p. 561-567.

Rosser, J. C., L. E. Rosser, and R. S. Savalgi, 1997, Skill acquisition and assessment for laparoscopic surgery: Arch.Surg., v. 132, no. 2, p. 200-204.

Royston, C. M., M. R. Lansdown, and W. A. Brough, 1994, Teaching laparoscopic surgery: the need for guidelines: BMJ, v. 308, no. 6935, p. 1023-1025.

Sackier, J. M., G. Berci, and M. Paz-Partlow, 1991, A new training device for laparoscopic cholecystectomy: Surg.Endosc., v. 5, no. 3, p. 158-159.

Sarker, S. K., A. Chang, C. Vincent, and A. W. Darzi, 2005, Technical skills errors in laparoscopic cholecystectomy by expert surgeons: Surg Endosc., v. 19, no. 6, p. 832-835.

Sarker, S. K., R. Hutchinson, A. Chang, C. Vincent, and A. W. Darzi, 2006, Self-appraisal hierarchical task analysis of laparoscopic surgery performed by expert surgeons: Surg.Endosc., v. 20, no. 4, p. 636-640.

Satava, R. M., 1999, Commentary.The need for metrics in surgical education: Surg.Endosc., v. 13, no. 11, p. 1082.

Satava, R. M., 2001b, Surgical education and surgical simulation: World J.Surg., v. 25, no. 11, p. 1484-1489.

Satava, R. M., 2001a, Accomplishments and challenges of surgical simulation: Surg Endosc., v. 15, no. 3, p. 232-241.

Schauer, P., S. Ikramuddin, G. Hamad, and W. Gourash, 2003, The learning curve for laparoscopic Roux-en-Y gastric bypass is 100 cases: Surg Endosc., v. 17, no. 2, p. 212-215.

Schijven, M., and J. Jakimowicz, 2002, Face-, expert, and referent validity of the Xitact LS500 laparoscopy simulator: Surg.Endosc., v. 16, no. 12, p. 1764-1770.

Schijven, M. P., and J. J. Jakimowicz, 2003, Introducing the Xitact LS500 laparoscopy simulator: toward a revolution in surgical education: Surg.Technol.Int., v. 11, p. 32-36.

Schirmer, B. D., S. B. Edge, J. Dix, and A. D. Miller, 1992, Incorporation of laparoscopy into a surgical endoscopy training program: Am.J.Surg., v. 163, no. 1, p. 46-50.

Scott, D. J. et al., 2000a, Laparoscopic training on bench models: better and more cost effective than operating room experience?: J.Am.Coll.Surg., v. 191, no. 3, p. 272-283.

Scott, D. J., R. V. Rege, P. C. Bergen, W. A. Guo, R. Laycock, S. T. Tesfay, R. J. Valentine, and D. B. Jones, 2000b, Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation: J.Laparoendosc.Adv.Surg Tech.A, v. 10, no. 4, p. 183-190.

Sedlack, R. E., and J. C. Kolars, 2004, Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study: Am.J.Gastroenterol., v. 99, no. 1, p. 33-37.

Seymour, N. E., A. G. Gallagher, S. A. Roman, M. K. O'Brien, D. K. Andersen, and R. M. Satava, 2004, Analysis of errors in laparoscopic surgical procedures: Surg.Endosc., v. 18, no. 4, p. 592-595.

Seymour, N. E., A. G. Gallagher, S. A. Roman, M. K. O'Brien, V. K. Bansal, D. K. Andersen, and R. M. Satava, 2002, Virtual reality training improves operating room performance: results of a randomized, double-blinded study: Ann.Surg., v. 236, no. 4, p. 458-463.

Seymour, N. E., and J. S. Rotnes, 2006, Challenges to the development of complex virtual reality surgical simulations: Surg.Endosc., v. 20, no. 11, p. 1774-1777.

Sharp, R. J., and J. Wellwood, 1996, The training implications of day-case surgery: Br.J.Hosp.Med., v. 55, no. 8, p. 472-475.

Sherman, V., L. S. Feldman, D. Stanbridge, R. Kazmi, and G. M. Fried, 2005, Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator: Surg Endosc., v. 19, no. 5, p. 678-682.

Shikora, S. A., J. J. Kim, M. E. Tarnoff, E. Raskin, and R. Shore, 2005, Laparoscopic Roux-en-Y gastric bypass: results and learning curve of a high-volume academic program: Arch.Surg, v. 140, no. 4, p. 362-367.

Silen, W., 2001, Crisis in surgical education: J.Am.Coll.Surg., v. 193, no. 5, p. 514-515.

Sjostrom, L. et al., 2004, Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery: N.Engl.J Med, v. 351, no. 26, p. 2683-2693.

Smith, C. D., T. M. Farrell, S. S. McNatt, and R. E. Metreveli, 2001, Assessing laparoscopic manipulative skills: Am.J.Surg., v. 181, no. 6, p. 547-550.

Smith, R., 1998, All changed, changed utterly. British medicine will be transformed by the Bristol case: BMJ, v. 316, no. 7149, p. 1917-1918.

Smith, S. G., J. Torkington, T. J. Brown, N. J. Taffinder, and A. Darzi, 2002, Motion analysis: Surg.Endosc., v. 16, no. 4, p. 640-645.

Soanes, C., and A. Stevenson, 2004,Soanes, C., and A. Stevenson Concise Oxford English Dictionary: Oxford, Oxford University Press.

Society of American Gastrointestinal Surgeons. Granting of Privileges for Laparoscopic General Surgery. Am.J.Surg. 161, 324-325. 1991.
Ref Type: Journal (Full)

Soper, N. J., J. A. Barteau, R. V. Clayman, S. W. Ashley, and D. L. Dunnegan, 1992, Comparison of early postoperative results for laparoscopic versus standard open cholecystectomy: Surg.Gynecol.Obstet., v. 174, no. 2, p. 114-118.

Spencer, F., 1978, Teaching and measuring surgical techniques: the technical evaluation of competence: Bull Am Coll Surg, v. 63, p. 9-12.

Stotter, A. T., A. J. Becket, J. P. Hansen, I. Capperauld, and H. A. Dudley, 1986, Simulation in surgical training using freeze dried material: Br.J.Surg., v. 73, no. 1, p. 52-54.

Street, K. N., 2004, Parent education classes: teaching junior paediatric doctors and student nurses about normal baby care: Med Teach., v. 26, no. 3, p. 273-276.

Sugerman, H. J., J. V. Starkey, and R. Birkenhauer, 1987, A randomized prospective trial of gastric bypass versus vertical banded gastroplasty for morbid obesity and their effects on sweets versus non-sweets eaters: Ann.Surg, v. 205, no. 6, p. 613-624.

Sutherland, L. M., P. F. Middleton, A. Anthony, J. Hamdorf, P. Cregan, D. Scott, and G. J. Maddern, 2006, Surgical simulation: a systematic review: Ann.Surg., v. 243, no. 3, p. 291-300.

Sweet, S. J., and I. J. Norman, 1995, The nurse-doctor relationship: a selective literature review: J Adv.Nurs., v. 22, no. 1, p. 165-170.

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

Taffinder,NJ, S G Smith, Mair J, Russell RCG, A Darzi. Can a computer measure surgical precision? Reliability, validity and feasibility of the ICSAD. Surg.Endosc. 13[S1], 81. 1999.
Ref Type: Abstract

Tang, B., G. B. Hanna, N. M. Bax, and A. Cuschieri, 2004a, Analysis of technical surgical errors during initial experience of laparoscopic pyloromyotomy by a group of Dutch pediatric surgeons: Surg.Endosc., v. 18, no. 12, p. 1716-1720.

Tang, B., G. B. Hanna, F. Carter, G. D. Adamson, J. P. Martindale, and A. Cuschieri, 2006, Competence assessment of laparoscopic operative and cognitive skills: Objective Structured Clinical Examination (OSCE) or Observational Clinical Human Reliability Assessment (OCHRA): World J.Surg., v. 30, no. 4, p. 527-534.

Tang, B., G. B. Hanna, and A. Cuschieri, 2005, Analysis of errors enacted by surgical trainees during skills training courses: Surgery, v. 138, no. 1, p. 14-20.

Tang, B., G. B. Hanna, P. Joice, and A. Cuschieri, 2004b, Identification and categorization of technical errors by Observational Clinical Human Reliability Assessment (OCHRA) during laparoscopic cholecystectomy: Arch.Surg., v. 139, no. 11, p. 1215-1220.

Taylor, H. L., D. A. Talleur, T. W. Emanuel, Jr., and E. M. Rantanen, 2005, Transfer of training effectiveness of a flight training device (FTD): Proceedings of the 13th International Symposium on Aviation Psychology, p. 1-4.

Tekkis, P. P., A. J. Senagore, C. P. Delaney, and V. W. Fazio, 2005, Evaluation of the learning curve in laparoscopic colorectal surgery: comparison of right-sided and left-sided resections: Ann.Surg., v. 242, no. 1, p. 83-91.

Tompkins, R. K., 1990, Laparoscopic cholecystectomy. Threat or opportunity?: Arch.Surg., v. 125, no. 10, p. 1245.

Torkington, J., S. G. Smith, B. I. Rees, and A. Darzi, 2001, Skill transfer from virtual reality to a real laparoscopic task: Surg.Endosc., v. 15, no. 10, p. 1076-1079.

Torkington, J., S. G. Smith, B. I. Rees, and A. Darzi, 2000, The role of simulation in surgical training: Ann.R.Coll.Surg Engl., v. 82, no. 2, p. 88-94.

Undre, S., A. N. Healey, A. Darzi, and C. A. Vincent, 2006, Observational assessment of surgical teamwork: a feasibility study: World J.Surg., v. 30, no. 10, p. 1774-1783.

Urbach, D. R., and N. N. Baxter, 2004, Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data: BMJ, v. 328, no. 7442, p. 737-740.

Van Rij, A. M., J. R. McDonald, R. A. Pettigrew, M. J. Putterill, C. K. Reddy, and J. J. Wright, 1995, Cusum as an aid to early assessment of the surgical trainee: Br.J.Surg., v. 82, no. 11, p. 1500-1503.

Van Sickle, K. R., D. A. McClusky, III, A. G. Gallagher, and C. D. Smith, 2005, Construct validation of the ProMIS simulator using a novel laparoscopic suturing task: Surg.Endosc., v. 19, no. 9, p. 1227-1231.

Vander Velpen, G. C., S. M. Shimi, and A. Cuschieri, 1993, Outcome after cholecystectomy for symptomatic gall stone disease and effect of surgical access: laparoscopic v open approach: Gut, v. 34, no. 10, p. 1448-1451.

Verdaasdonk, E. G., L. P. Stassen, L. J. Monteny, and J. Dankelman, 2006, Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO: Surg.Endosc., v. 20, no. 3, p. 511-518.

Vincent, C., K. Moorthy, S. K. Sarker, A. Chang, and A. W. Darzi, 2004, Systems approaches to surgical quality and safety: from concept to measurement: Ann.Surg, v. 239, no. 4, p. 475-482.

Vitale, G. C., D. Collet, G. M. Larson, W. G. Cheadle, F. B. Miller, and J. Perissat, 1991, Interruption of professional and home activity after laparoscopic cholecystectomy among French and American patients: Am.J.Surg., v. 161, no. 3, p. 396-398.

Voyles, C. R., A. B. Petro, A. L. Meena, A. J. Haick, and A. M. Koury, 1991, A practical approach to laparoscopic cholecystectomy: Am.J.Surg., v. 161, no. 3, p. 365-370.

Ward, M., H. MacRae, C. Schlachta, J. Mamazza, E. Poulin, R. Reznick, and G. Regehr, 2003, Resident self-assessment of operative performance: Am.J.Surg., v. 185, no. 6, p. 521-524.

Warf, B. C., M. B. Donnelly, R. W. Schwartz, and D. A. Sloan, 1999, Interpreting the judgment of surgical faculty regarding resident competence: J.Surg.Res., v. 86, no. 1, p. 29-35.

Weber, M., M. K. Muller, T. Bucher, S. Wildi, D. Dindo, F. Horber, R. Hauser, and P. A. Clavien, 2004, Laparoscopic gastric bypass is superior to laparoscopic gastric banding for treatment of morbid obesity: Ann.Surg, v. 240, no. 6, p. 975-982.

Weigelt, J. A., 2003, Medical education crisis: not just an issue of work hours: Arch.Surg., v. 138, no. 9, p. 1027-1028.

Wetzel, C. M., R. L. Kneebone, M. Woloshynowych, D. Nestel, K. Moorthy, J. Kidd, and A. Darzi, 2006, The effects of stress on surgical performance: Am.J.Surg., v. 191, no. 1, p. 5-10.

Wilson, M. S., A. Middlebrook, C. Sutton, R. Stone, and R. F. McCloy, 1997, MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance: Ann.R.Coll.Surg.Engl., v. 79, no. 6, p. 403-404.

Winckel, C. P., R. K. Reznick, R. Cohen, and B. Taylor, 1994, Reliability and construct validity of a structured technical skills assessment form: Am.J.Surg., v. 167, no. 4, p. 423-427.

Wittgrove, A. C., G. W. Clark, and L. J. Tremblay, 1994, Laparoscopic Gastric Bypass, Roux-en-Y: Preliminary Report of Five Cases: Obes.Surg, v. 4, no. 4, p. 353-357.

Wood, R. Y., 1994, Use of the nursing simulation laboratory in reentry programs: an innovative setting for updating clinical skills: J Contin.Educ.Nurs., v. 25, no. 1, p. 28-31.

Woodrum, D. T., P. B. Andreatta, R. K. Yellamanchilli, L. Feryus, P. G. Gauger, and R. M. Minter, 2006, Construct validity of the LapSim laparoscopic surgical simulator: Am.J.Surg., v. 191, no. 1, p. 28-32.

Working group on specialist medical training. Hospital doctors: training for the future. 1993. London, Department of Health.
Ref Type: Report

Yanovski, S. Z., and J. A. Yanovski, 2002, Obesity: N.Engl.J Med, v. 346, no. 8, p. 591-602.

Youngblood, P. L., S. Srivastava, M. Curet, W. L. Heinrichs, P. Dev, and S. M. Wren, 2005, Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance: J.Am.Coll.Surg., v. 200, no. 4, p. 546-551.

Zeger, S. L., K. Y. Liang, and P. S. Albert, 1988, Models for longitudinal data: a generalized estimating equation approach: Biometrics, v. 44, no. 4, p. 1049-1060.

Zucker, K. A., R. W. Bailey, T. R. Gadacz, and A. L. Imbembo, 1991, Laparoscopic guided cholecystectomy: Am.J.Surg., v. 161, no. 1, p. 36-42.