Acquiring minimally invasive surgical skills

Hiemstra, E.

ACQUIRING MINIMALLY INVASIVE SURGICAL SKILLS

ISBN: 978-94-6182-060-0

Cover illustration: Marjon van Tongeren, www.marjonvantongeren.com

Illustration design: Suzanne Hiemstra–Van Mastrigt

Lay-Out and Printing: Off Page, Amsterdam

Financial support for the publication of this thesis was provided by:

A Hiemstra-Timmenga, NVEC, Simendo BV, Maatschap Gynaecologie Haga Ziekenhuis, DSSH, BMA BV (Mosos), Memedis Pharma BV, Convidien Nederland BV, Johnson & Johnson Medical BV, Olympus Nederland BV, Janssen Cilag BV, Medical Dynamics BV.

ACQUIRING MINIMALLY INVASIVE SURGICAL SKILLS

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit van Leiden
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden
volgens besluit van het College van Promoties
te verdedigen op donderdag 26 januari 2012
klokke 15.00 uur

door

Ellen Hiemstra
geboren te Zwolle in 1979

Promotiecommissie

Promotores	Prof. dr. F.W. Jansen
	Mw. prof. dr. ir. J. Dankelman, Technische Universiteit, Delft
Overige leden	Prof. dr. J.F. Hamming
	Mw. dr. M.P. Schijven, Academisch Medisch Centrum, Amsterdam
	Prof. dr. J.B.M.Z. Trimbos

Aan mijn ouders

CONTENTS

Chapter 1 General Introduction 9

PART I Outside of the Operating Room

Chapter 2 Skills training in minimally invasive surgery in Dutch Obstetrics and Gynaecology Residency Curriculum 15

Chapter 3 Virtual Reality in Laparoscopic Skills Training: Is Haptic Feedback replaceable? 21

Chapter 4 Optimizing laparoscopic skills training: Does a fixed camera compromise depth perception? 29

Chapter 5 Intracorporeal suturing: economy of movements in a box trainer model 37

Chapter 6 Retention of basic laparoscopic skills after a structured training program 47

Chapter 7 Grading surgical skills curricula and training facilities for minimally invasive surgery 55

PART II In the Operating Room

Chapter 8 The value of an objective assessment tool in the operating room 63

Chapter 9 Implementation of OSATS in the Residency Program: a benchmark study 73

Chapter 10 Are minimally invasive procedures harder to acquire than conventional surgical procedures? 81

Chapter 11 General discussion 89

Chapter 12 Conclusions and Recommendations 95

Chapter 13 Summary / Samenvatting 99

Chapter 14 **Addendum**

Literature 111

Author affiliations 119

About the author 121

List of publications 123

Dankwoord 125

CHAPTER 1
GENERAL INTRODUCTION

Minimally invasive surgery (MIS) developed due to technological advances in instrumentation along with an appreciation that avoidance of laparotomy may confer advantages for patient recovery such as reduced post-operative pain, shorter hospitalisation, more rapid return to normal activities, and improved cosmetic results.[Darzi et al., 1999] Still, MIS has some different surgical features in comparison with laparotomy. In the first place, the depth perception of the surgeon is reduced because the operation field has to be interpreted from a two dimensional (2D) screen.[Munz et al., 2004] Furthermore, long instruments are inserted through the abdominal wall during laparoscopy. This creates counter intuitive movements with a limited range of motion and results in an distorted hand eye coordination.[Gallagher et al., 1998; Pearson et al., 2002] In the third place, haptic feedback is diminished, because there is no direct contact between the surgeon's gloved hands and the tissue [Bholat et al., 1999]. Finally, camera instability may increase fatigue.[Heemskerk et al., 2006] As a consequence, a surgeon who performs MIS is faced with the challenge to master a different set of technical surgical skills compared to performing a conventional procedure.

Despite the advantages of MIS for patient recovery, this new surgical technique has not been adopted without any trouble. The initial implementation of the laparoscopic cholecystectomy progressed rapidly and has led to an alarming number of significant complications due to inadequately trained and skilled surgeons.[Forde, 1993] These concerns remained after its initial adoption phase, as illustrated by the report published by the Dutch Health Care Inspectorate, entitled "Risks of minimally invasive surgery underestimated".[IGZ 2007] The Inspectorate stated that the actions taken to prevent incidents in MIS were insufficient. Specifically, improvement of the training of MIS skills was demanded, combined with the setting of a certain level of basic endoscopic skills prior to operating on real patients.[Stassen et al., 2010]

Although the obligation for skills training was new with the publication of this report, the importance of basic MIS skills training outside of the OR has long been realized. In 1985, pelvi-trainers were already introduced by Kurt Semm to learn 'how to operate mono-and binocular' and to 'handle the grips of the instruments' (Figure 1).

Unfortunately, no broad implementation of these boxes occurred. However, the following arguments support skills training outside of the OR prior to patient exposure. Firstly, there are ethical concerns about teaching basic skills on a patient, when alternatives are readily available.

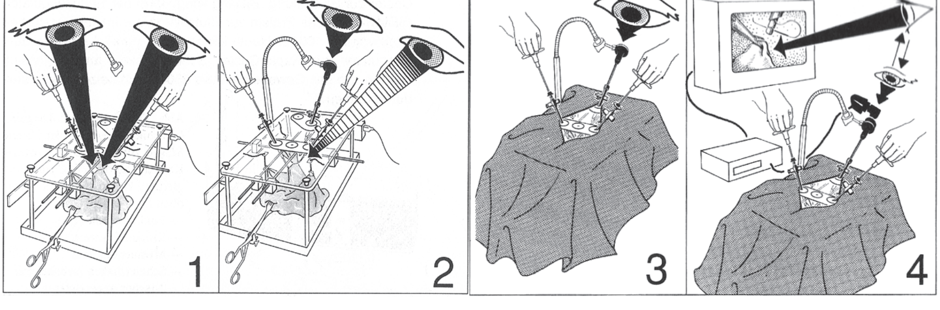


Figure 1. pelvi-trainer as designed by Kurt Semm.

Skills acquired on box trainers [Scott et al., 2000] and virtual reality (VR) trainers [Grantcharov et al., 2004; Seymour et al., 2002] are transferable to surgery on real patients. Moreover, simulator training might bypass the early learning curve, which is known to be associated with an increased rate of complications. [Southern Surgeons Club, 1991] A second argument to support that the OR should not be the predominant learning environment is that surgeons are pressed to be more efficient in the OR due to increasing financial constraints. Thirdly, teaching hospitals are increasingly populated by patients with more serious and complex surgical problems that demand the skills of expert surgeons working at maximum efficiency.[Blanchard et al., 2004; Brolmann et al., 2001] Finally, working hours restrictions leave residents with less opportunities to perform surgical procedures on living humans. Teaching fundamental MIS skills outside of the OR is likely to improve the time trainees spend in the operating theatre because those who have acquired basic surgical skills can focus more thoroughly upon the anatomy, pathology and procedural aspects of actual surgery.[Korndorffer, Jr. et al., 2005c] Therefore, from a teaching perspective, it is more efficient to learn basic surgical skills prior to performing actual surgery.

Consequently, skills laboratories have been set up worldwide in order to train and assess MIS skills outside of the OR. By now, most teaching hospitals have training facilities, or at least access to it elsewhere. However, no guidelines or standards exist yet how to design and use such facilities. This parallels the finding of the Dutch Health Inspectorate that there is no uniformity in MIS training, both in general and between endoscopic professionals (e.g. surgeons, gynaecologists, urologists).[Stassen et al., 2010] In fact, the development of training facilities is often based upon the funder's personal preferences and on the money available, rather than upon scientific evidence of the value for the process to acquire surgical skills. Even, it has been stated that one of the greatest errors in setting up a surgical skills lab, is to purchase the equipment first and then to design a curriculum around it.[MacRae et al., 2008]

Up until now, most research has focused on the available trainers for MIS surgery. The two inanimate simulators are a box trainer and a VR trainer. Within these two categories, many types of simulators have been developed with even more exercises, varying from relatively simple tasks to entire procedures.[Hammoud et al., 2008] Mounting validation studies have been conducted on new exercises in the box or VR trainers. [Kolkman et al., 2008; Schreuder et al., 2011] Validity addresses the concept of whether the test is actually measuring what it was intended to measure.[Feldman et al., 2004b] For example, does the simulator discriminate surgeons of different skill levels, and does the exercise resemble an actual surgical situation? Validity is a prerequisite before exercises are employed in a MIS training program. However, there is no consensus about the optimal type of trainer.[Stefanidis et al., 2009b] Additionally, it is unknown which metrics should be applied for simulator training and assessment purposes, with existing measures varying from simple (time) to the more complex (motion analysis). [Hammoud et al., 2008] Although, evidence is emerging that training should be mandatory,[Hammoud et al., 2008; Kolkman et al., 2007b; Stefanidis et al., 2008] formal curriculum development is lagging behind.

In summary, more evidence is required to identify training resources, exercises and programs which confer the best outcomes in terms of acquiring proficiency in predefined training objectives. The first part of this thesis addresses the questions raised above, and thereby aims at providing a more solid, scientific basis for the design and the use of MIS training facilities outside of the OR. Future investment should support those training facilities utilizing evidence based training.

Despite the importance of skills training facilities outside of the OR, the real craft of surgery is obviously transmitted in the OR. Moreover, the decision-making processes and sequels of errors possibly leading to complications cannot be trained with the use of inanimate training models and only partly with the use of animate ones.[Schijven et al., 2004] In fact, in the 100-year-old Halstedian teaching model, the OR was the only place where residents acquired their technical surgical skills. The adage "see one, do one, teach one" was the motto of the surgical training program. Techniques and views were simply handed down from the senior surgeon to the resident until he or she was believed capable of performing surgery independently. The evaluation was coloured by subjectivity.[Darzi et al., 1999] The opinion of the supervising surgeon was practically the only standard that had to be met.[Schijven et al., 2008] This method, also called the *apprenticeship model*, has produced generations of fine technical surgeons.[Haluck et al., 2000] However, it may no longer be optimal with accelerating changes in the health care system: Authority and public demand a safer and more transparent health care system, rather than automatically accepting the proficiency of surgeons. Additionally, specialty training is moving towards more competency based outcome measures rather than being solely based on the training length. To achieve this, more objective external assessments are needed for accurate appraisal in the challenging area of surgical proficiency.[Aggarwal et al., 2004]

Examples of assessment tools for surgical skills are the OSATS (Objective Structured Assessment of Technical Skills) [Martin et al., 1997] and the GOALS (Global Operative Assessment of Laparoscopic Skills) [Vassiliou et al., 2005]. OSATS was developed in Canada and was originally designed to measure technical surgical performance using six stations in a skills laboratory. [Martin et al., 1997] The six stations comprised of the excision of a skin lesion, hand sewn bowel anastomosis, stapled bowel anastomosis, insertion of a T-tube, abdominal wall closure and control of inferior vena cava haemorrhage. The authors established the validity, reliability and feasibility of the general global rating scale of the OSATS for these six tasks. Subsequently, the value of the OSATS has been proven for large scale implementation in obstetrics and gynaecology residency programs, but again it only focused on its use in a laboratory setting. [Goff et al., 2005] However, in the Netherlands this method has been introduced for evaluating surgical skills during real procedures in the OR. This introduction took place in absence of data on the validity of the OSATS in the real surgical setting. Therefore, the second part of this thesis focuses on whether evidence is present to use the OSATS as an intraoperative assessment tool either in conventional and MIS procedures.

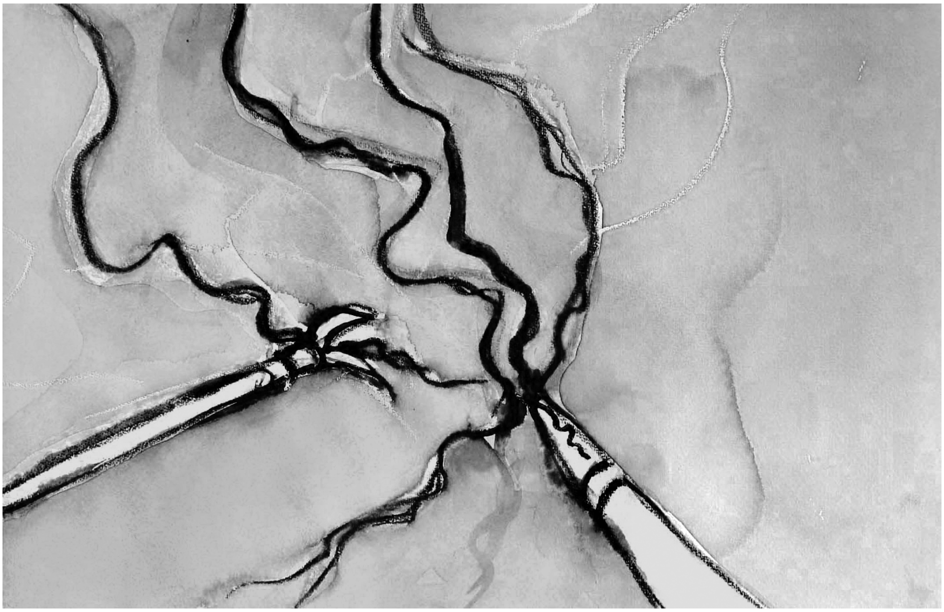
OUTLINE OF THE THESIS

As an introduction to the organisation of MIS skills training, **Chapter 2** describes a mandatory nationwide surgical skills course in the Netherlands, with a critical discussion of the current system. As mentioned before, there is no consensus which trainer model should be chosen, especially with constant improvements in trainer models. In VR trainers, the addition of kinematic interaction between laparoscopic instruments and objects is a possible solution to compensate for the lack of haptic feedback. In **chapter 3** we determined whether or not this interaction can replace the haptic feedback that is naturally present in box trainers. A comparison between box and VR trainers is made with respect to acquiring tissue handling skills. Furthermore, both fixed and navigated camera setups are available during simulator training. A navigated camera offers theoretical advantages for the depth perception of the surgeon and allows the practice of navigation skills, whereas a fixed setup allows solitary training. The effect of camera setup on surgical performance is yet unknown. Therefore, three different camera setups are compared in **chapter 4**. As a next step after the choice for a training model**, chapter 5** focuses on the metrics used for training and assessment in a box trainer. In addition to time, three movement analysis parameters are validated for the clinically important knot tying task, by using a tracking device. Regarding the organisation of a skills curriculum, we investigated in **chapter 6** whether the skills acquired during five validated box trainer tasks remain after one year. Skills laboratories have been set up in teaching hospitals all over the world for the training and assessment of MIS skills. However this has been done in the absence of generally accepted standards as to what a MIS skills laboratory should look like and how the training should be conducted. In **chapter 7** an international and consensus based set of quality criteria is developed for a MIS training skills laboratory, including the design of the laboratory and the training curriculum.

Although the OSATS have proved to be valid, feasible and reliable for the use in a laboratory setting, its value for intraoperative use still needs to be established. **Chapter 8** evaluates the validity of this tool for intraoperative use. In addition, more issues relevant to the implementation of the OSATS as an intraoperative assessment tool are studied in **chapter 9**. Firstly, it is determined at which OSATS score a resident is able to perform a certain procedure autonomously. Secondly, the concurrence in the assessment by supervisor and resident is established as a measure of its reliability. Thirdly, the feasibility is investigated by a survey among residents and staff confronted with the tool in daily practice. In **chapter 10**, the OSATS is used as a reference to answer the question as to whether MIS procedures are harder to acquire for the current generation of residents? This answer is found by comparing residents' learning curve for MIS procedures with the curve for conventional surgical procedures.

In **chapter 11** the research results are outlined in a general discussion. This is followed by conclusions and recommendations in **chapter 12**. Finally, in **chapter 13** this thesis is summarized in English and Dutch.

CHAPTER 2
SKILLS TRAINING IN MINIMALLY INVASIVE SURGERY IN DUTCH OBSTETRICS AND GYNAECOLOGY RESIDENCY CURRICULUM



Ellen Hiemstra
Wendela Kolkman
Frank Willem Jansen

Adapted from Gynecol Surg 2008; 5: 321-5

INTRODUCTION

2

Minimally invasive surgery (MIS) has evolved into a major surgical approach to treat a variety of gynaecological disorders. This approach has considerable benefits for patients, such as a reduced morbidity, a shorter hospitalization, better cosmetic results, and an earlier return to normal activity.[Darzi et al., 2002]

However, acquiring MIS skills is more challenging than acquiring the skills necessary to perform conventional open surgical procedures. MIS poses specific demands on the surgeon. During MIS the three-dimensional operating field has to be interpreted from a two-dimensional monitor display in which depth perception is altered. In addition, a surgeon has to manipulate long surgical instruments with diminished tactile feedback and fewer degrees of freedom, while adapting to the fulcrum effect.[Gallagher et al., 1998; Munz et al., 2004]

Apart from the complexity of acquiring MIS skills, a residency curriculum has to deal with smaller case volumes in the operating room (OR). This is due to a decrease in resident working hours and a declining trend in major gynaecological surgical procedures in general.[Blanchard et al., 2004; Brolmann et al., 2001] The smaller case volumes, combined with issues such as quality control, patient safety, efficiency and cost-effectiveness have led to an increasing interest in simulator training facilities outside the OR.[Feldman et al., 2004b; Munz et al., 2004] Simulator training aims at progression along the learning curve by repetitive training of surgical skills with a lack a potential burden to patients in a pressure free environment.[Munz et al., 2004]

With respect to MIS training, the implementation into residency programs is shown to be troublesome.[Loh et al., 2002; Navez et al., 1999; Nussbaum, 2002] Even though basic laparoscopic procedures have well been incorporated in residency, more advanced procedures are not.[Brolmann et al., 2001; Kolkman et al., 2005] Lack of adequate training during residency influences the subsequent use of a specific technique and ultimately may restrict the implementation of MIS in daily practice after completion of residency training.[Kolkman et al., 2006; Shay et al., 2002]

In this report we present the organization of MIS skills training in the Dutch obstetrics and gynaecology residency curriculum which has continuously been evaluated and improved over the past 15 years.

SURGICAL SKILLS IN THE DUTCH RESIDENCY CURRICULUM

The obstetrics and gynaecology residency program lasts six years in the Netherlands. A basic surgical skill course, named the Cobra-alpha course, was incorporated in the curriculum in 1992. It has been evaluated and improved ever since. Attendance to this course was made compulsory for residents obstetrics and gynaecology in 1997, and they had to attend it during postgraduate year (PGY) 1 or 2. resident. One third of this two-day course is spent on theory, while the complementary two thirds are spent on hands-on training. The first day focuses on basic technical skills, like instrument handling and knot tying, for conventional surgery, while the second day concerns the basic skills required for MIS which is subdivided into laparoscopy and hysteroscopy. Three handbooks, focusing on the basics of surgery, hysteroscopy and laparoscopy, are used for study purposes and have been written for this course.[Jansen F.W. et al., 2008; Jansen et al., 2006; Trimbos J.B., 2007]

The goal of the hands-on training in MIS during the Cobra-alpha course is to provide an introduction to simulator training for laparoscopic and hysteroscopic skills. Additionally, residents need to expand the acquired skills on simulators and have these skills evaluated by a mentor or MIS expert in their own clinic. Necessarily, time for training and evaluation has to be scheduled into the busy clinical practice of the residency program.

A range of simulators is available for the hands-on training. Inanimate box trainers are used to practice basic laparoscopic skills like hand-eye coordination, adaptation to the lack of depth perception and camera holding. The construct validity is established for five of the available exercises in the box trainers.[Kolkman et al., 2008] These five exercises are placing a pipe cleaner through four small circles, stretching a rubber band around 16 nails, placing 13 beads in a letter 'B', cutting a marked circle from a rubber glove and intra-corporeal knot tying. The laparoscopic box trainer exercises are presented in figure 1. With regard to hysteroscopic simulators, vegetable models are available like pumpkins and red peppers.[Kingston et al., 2004] Furthermore, a chicken meat simulates endometrium in a water filled box and a porcine bladder simulates a uterus. A selection of hysteroscopic exercises is presented in figure 2. Basic hysteroscopic skills such as camera holding, instrument handling, safe use of energy sources and distension medium are trained. Besides, some procedures like diagnostic hysteroscopy, endometrium resection and resection of polyps or myomas are simulated.

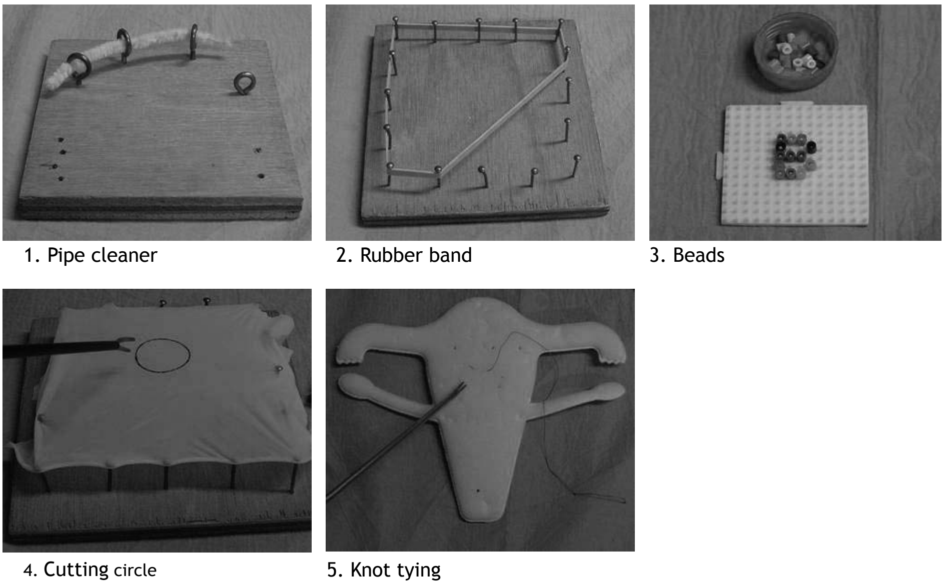
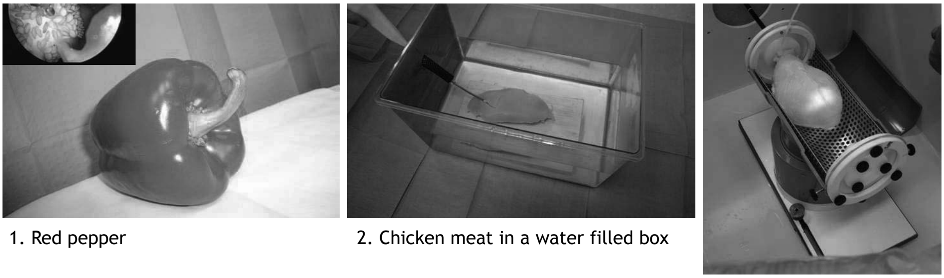


Figure 1. Laparoscopic training exercises.



1. Red pepper 2. Chicken meat in a water filled box 3. Porcine bladder

Figure 2. Hysteroscopic training exercises.

Prior to the start of hands-on training, the exercises are introduced and explained with the aid of audio-visual demonstration. Afterwards, the participants go through a rotation of simulators. The surgical performance is assessed by calculating a score that rewards precision and speed. In the validated exercises, the calculated individual scores are compared to a previously established performance standard.[Kolkman et al., 2008] Training on the laparoscopic and hysteroscopic simulators is intensively supervised by experts in MIS. Regarding the number of participants attending the course, which varies from 32 to 36, each simulator is used by two or three residents and is supervised by one supervising expert.

In addition to the mandatory Cobra-alpha course which is mainly focused on basic skills, residents can apply to two advanced courses in MIS, a laparoscopy course and a hysteroscopy course. These courses can be attended on a voluntary basis. The advanced courses are more procedure orientated than the Cobra-alpha course. In spite of using simulators, life surgery is used for teaching purposes. Procedure specific courses can further enhance skills and knowledge, like a sacrocolpopexy course and a course regarding laparoscopic adnex surgery. Besides, a variety of (inter)national congresses focuses on MIS are organized.

The mandatory Cobra-alpha course, advanced MIS courses and congresses form the training structure in the Dutch residency curriculum, combined with simulator training in the teaching hospitals during clinical rotation.

DISCUSSION

The Dutch obstetrics and gynaecology residency curriculum has a clear structure regarding the training of MIS skills. A mandatory basic surgical skills course is established for residency training which is nationwide accepted and has a broad Dutch faculty. Intentionally, the course has to be attended during PGY 1 or PGY 2. Additionally, residents may attend advanced courses and congresses focusing on laparoscopy and hysteroscopy. This structure enhances the implementation of basic MIS skills training into the residency curriculum.

Basic MIS skills can be trained on simulators. Simulators have shown great potential for training and objectively assessing laparoscopic skills.[Lentz et al., 2001; Scott et al., 2001] The skills acquired are transferable to real operative procedures[Anastakis et al., 1999; Hyltander et al., 2002; Seymour et al., 2002] and skills training is shown to decrease patient complications.[Cadeddu et al., 2001] For every resident there is a learning curve to achieve proficiency in performing MIS. Presumably, acquiring basic MIS skills by simulator training leads to progression along the first part of this learning curve resulting in better prepared residents for the actual surgery. After achievement of the basic skills, more attention can be paid to the specific procedure during surgery on real patients. With the growing evidence of valuable aspects of MIS simulator training, we feel there is no excuse for depriving residents of this training.

The nationwide basic surgical skills course provides an introduction in simulator training for acquiring MIS skills. However, distributed practice is superior above massed practice, which is provided during a two-day course, for actually achieving these skills.[Moulton et al., 2006; Verdaasdonk et al., 2007b] Consequently, MIS skills can only be acquired if residents continue simulator training and evaluation in their own clinic. A first precondition for this continuance of training is the presence of simulator facilities in every cluster of teaching hospitals. A second precondition is that residents really do use these facilities. The first precondition is partially met. All 46 Dutch teaching hospitals are grouped in eight clusters and simulator training is offered in at least one teaching hospital of each cluster. However, the equipment varies widely among these hospitals. The advantage of training on the simulators used during the Cobra-alpha course is that these are easily fabricated and inexpensive. Besides, the exercises for the laparoscopic box trainer have been validated and a performance standard has been established. Regarding the second precondition, unfortunately only one third of residents actually train on a simulator if training is offered on a voluntary basis.[Kolkman et al., 2005] The fact that most residents do not voluntarily train is in contradiction with the residents' opinion that simulator training is an important addition to their residency program.[Kolkman et al., 2005] Hence, formal mandatory MIS training is urgently needed in every training hospital, which has to be scheduled in the busy practice of the residency program.

In spite of structured training, proper evaluation of skills contributes to the learning effect.[Reznick et al., 1997] However, the majority of residents' surgical skills are evaluated informally and in a non-standardized fashion. There is a growing need for objective assessment tools. An example of such a tool is the Objective Structures Assessment of Technical Skills (OSATS). This evaluation method consists of a global rating scale and has proven high reliability and construct validity for simulators.[Goff et al., 2002; Reznick et al., 1997]

Regarding surgical competence, the requirements essential for certification in obstetrics and gynaecology are clearly defined in the Netherlands. These requirements are set on a total number of each procedure a resident minimally has to perform. Additionally, the number performed on competence level 4 is established. Level 4 is defined "able to perform without supervision" on a 1 to 5 global rating scale (Table 1). The target numbers for the laparoscopic and hysteroscopic procedures are expressed in Table 2.[NVOG-HOOG 2005] Although numbers of procedures are easily quantifiable, total numbers do not represent the actual competence of a resident due to individual difference in learning curves.[Park et al., 2002a] Assessing a residents surgical skills and comparing these skills to an established performance standard would be a more suitable than counting the number of procedures. In this way, the individual training demands can be met. This emphasizes on one hand the importance of objective assessment tools for evaluation of surgical skills and to set a performance standard. On the other hand simulator training can fulfil the individual training demands, as a source of unlimited training while the training possibilities on real patients in the OR are scarcer. Ultimately, every resident should be able to achieve the predetermined level of skills at the end of residency.

Table 1. Global rating scale for level of competence.

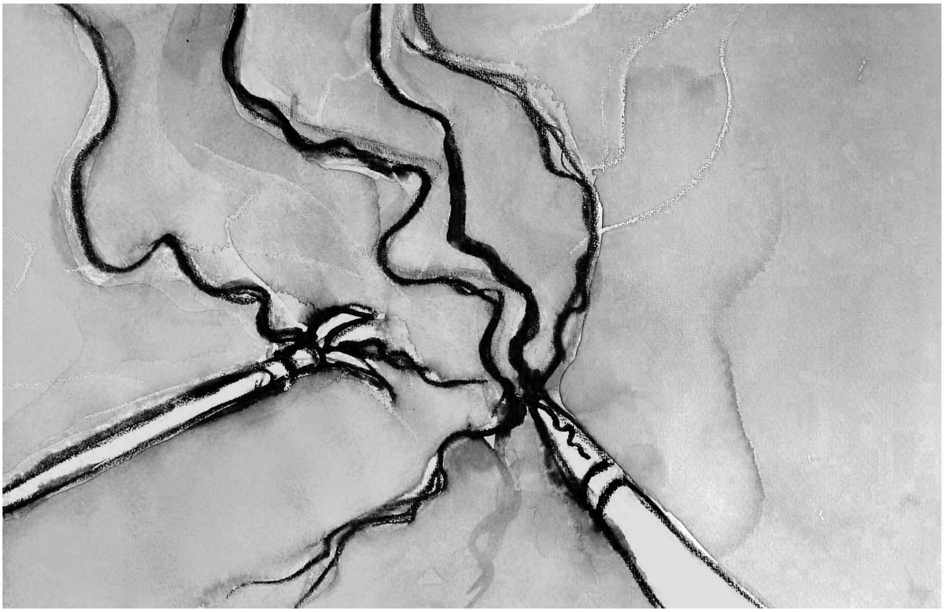
Level	Definition
1	Has theoretical knowledge
2	Is able to perform under strict supervision
3	Is able to perform under limited supervision
4	Is able to perform without supervision
5	Is able to supervise and educate others

Table 2. Target numbers of MIS procedures required for certification

Procedure	Target number (Total)	Target number performed on competence level 4
Laparoscopic surgery		
Diagnostic laparoscopy / sterilization	50	10
Minor adhesiolysis	10	not applicable
Salpingectomy / salpingotomy (inclusive EP)	20	5
Cystectomy	10	not applicable
Hysteroscopy		
Diagnostic hysteroscopy	40	10
Resection polyps	10	5
Resection myomas type 0-I	10	not applicable
Resection myomas type II	10	not applicable

Although some adaptations have to be made to incorporate continued training and evaluation in daily practice, a uniform introduction to MIS training on simulators is guaranteed for every resident in the Netherlands by a mandatory basic skills course, while advanced courses and congresses provide possibilities for enhanced education. Hopefully, this will facilitate and accelerate the implementation of MIS techniques in the gynaecological surgical palette.

CHAPTER 3
VIRTUAL REALITY IN LAPAROSCOPIC SKILLS TRAINING: IS HAPTIC FEEDBACK REPLACEABLE?



Ellen Hiemstra
Elisabeth M. Terveer
Magdalena K. Chmarra
Jenny Dankelman
Frank Willem Jansen

Adapted from Minimally Invasive Therapy & Allied Techniques, 2011; 20: 79-84

3

INTRODUCTION

Surgeons traditionally rely on vision and touch to obtain information about the operation field. In minimally invasive surgery (MIS), like laparoscopy, these senses can only provide indirect information. Regarding vision, the operation field has to be interpreted from a two-dimensional projection of the endoscopic view.[Westebring-van der Putten EP et al., 2008] Regarding touch, the gloved surgeon's hand is in indirect contact with the tissue through laparoscopic instruments. The latter results in limited haptic (kinaesthetic and tactile) feedback.[Bholat et al., 1999] However, correct perception of the operation field is essential to guarantee efficient and safe tissue manipulation. Consequently, laparoscopic surgeons have to be capable of correctly interpreting indirect visual and haptic feedback.

The development of training facilities outside the operating room (OR) has taken a great leap. One of the explanations is that the apprenticeship model turned out to be insufficient for acquiring MIS skills.[Aggarwal et al., 2004] For training on inanimate models, numerous simulators have been introduced and validated.[Stefanidis et al., 2009b] Roughly, these simulators are divided into physical box trainers and computer-aided virtual reality (VR) trainers.[Dunkin et al., 2007] In box trainers, real laparoscopic instruments are used. A consequence of training with real laparoscopic instruments is that realistic haptic feedback is provided in box trainers. None of the VR trainers provides natural haptic feedback. Therefore, these devices are mainly focused on training hand eye-coordination.[Schijven et al., 2003]

Haptic feedback is considered necessary for tissue handling in laparoscopy. It is used to regulate force application, and thereby, avoid tissue damage.[Strom et al., 2006] Furthermore, it provides information on tissue texture, shape and consistency. Despite its clinical significance, little is known about the exact role of haptic feedback during simulator training. Only a few studies revealed its importance in the early training phase of skills acquisition.[Botden et al., 2008; Strom et al., 2006] Obviously, with respect to haptic feedback, box training models are superior to VR systems.

In response, attempts have been made to compensate the lack of haptic feedback in VR trainers by adding electromechanically transmitted information.[Westebring-van der Putten EP et al., 2008] This allows the trainee to "feel" an illusion of contact in the grip of the instrument. However, current technology is not yet able to provide it in a highly realistic manner.[Basdogan et al., 2004; Schijven & Jakimowicz, 2003; Westebring-van der Putten EP et al., 2008] Others tried to compensate for the lack of haptic feedback by using software that simulates real-time instrument tissue interactions based on instrument movements and imaginary physical properties of the objects in the virtual environment.[Basdogan et al., 2007] It is unknown whether tissue handling skills can be acquired using a VR trainer model equipped with this software. Therefore, the aim of this study is to determine whether (and to which extent) additional kinematic interaction in VR trainers can replace haptic feedback during laparoscopic skills training, by comparing the effect of box and VR training with different levels of kinematic interaction.

MATERIALS AND METHODS

This study was conducted at the skills laboratory of the Leiden University Medical Centre (LUMC) in the Netherlands from 2008 to 2009. The SIMENDO® VR trainer (Delltatech, Rotterdam, The Netherlands) was used for VR training setups. A physical box trainer (LUMC, Leiden) was used for the box trainer setups.

Study population

Novices (i.e. medical students in the preclinical phase of their studies) were recruited to the study by means of advertisement in the medical library of the LUMC. They participated on a voluntary basis. After enrolment, they completed a questionnaire providing demographic information (i.e. gender, hand dominancy, self-perceived dexterity, prior laparoscopic or simulator experience, and experience in computer gaming).

Study design

As a pre-test, all participants performed a validated[Kolkman et al., 2008] rubber band task in a box trainer. To fulfil this task, the rubber band first had to be put outside all 16 nails on the wooden board. Then, it had to be zigzagged around the nails, starting in the upper left corner. This task was chosen to simulate tissue handling during laparoscopic surgery, because it requires hand eye coordination as well as a proper application of forces.

After pre-testing, novices were randomly assigned to one of four training setups and a control group that received no training (Figure 1). In all training setups, which will be described in detail in the next section, participants performed an exercise to pile up three cylinders. Duration of the training was 20 minutes, the control group waited during that period. The duration of 20 minutes had been based on a pilot study in which we found that most of the short-term training effect was achieved within 20 minutes regarding piling up cylinders correctly in box and VR setups. The rubber band task was performed again as a post-test after training or waiting. The flowchart of the study is presented in figure 1.

Intervention – the training setups

In the *VR-I* setup, the cylinder task of the basic curriculum of SIMENDO® (SimSoft Basic 1.0 package) was used. Such a setup allows psychomotor skills training in a conventional VR environment. The curriculum has been validated and has shown to improve OR performance. [Verdaasdonk et al., 2007a] In the *VR-II* setup, the cylinder task of the new Simsoft Advanced 2.0 package of the SIMENDO® was used. The kinematic behaviour of the objects in the VR environment has been changed by adding object movements based on instrument's velocity and the physical properties of the objects (e.g. weight). Consequently, VR-II has a different kinematic instrument-object interaction based on calculated forces that are virtually applied. With these kinematic properties, it is, for example, determined whether a tower of cylinders will fall over when the table they are placed on tilts due to the virtual forces applied with a virtual laparoscopic grasper. The *Box-I* and the *Box-II* setups have been designed to be an equivalent of the VR-I setup and the VR-II setups, respectively. The only difference between these setups was that the table, on which the cylinders were placed, was fixed in Box-I, whereas in Box-II the

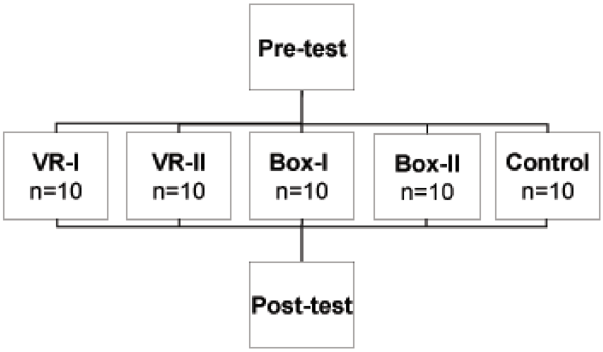


Figure 1. Study design. VR-I: set-up in conventional VR environment, VR-II: set-up with kinematic object interaction application, Box-I: box trainer equivalent of VR-I, Box-II: box trainer equivalent of VR-II. Control: no training. Participants were equally distributed to the five groups using randomization using the website www.randomization.com.

legs of the table were replaced by springs in order to allow the table to tilt. The endoscopic view of the four training setup is shown in figure 2.

The image of a fixed 0° scope was presented on the monitor in all training setups. Participants used two laparoscopic graspers, one in the right and one in the left hand. The dimensions of grasper of laparoscopic instrument, the cylinders, and the square table (on which cylinders were placed) were identical in each training setup. Consequently, the training varied with respect to the absence or presence of haptic feedback (i.e. the VR, and the box

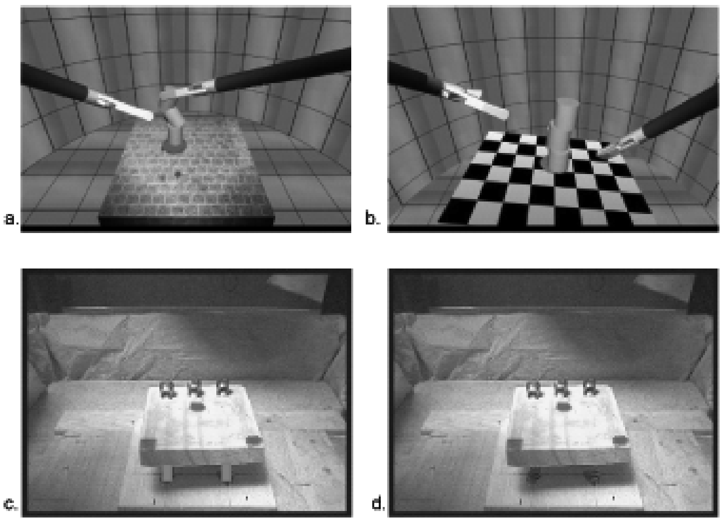


Figure 2. Four training setups. (a) VR-I: set-up in conventional VR environment, (b) VR-II: set-up with kinematic object interaction application, (c) Box-I: box trainer equivalent of VR-I, (d) Box-II: box trainer equivalent of VR-II.

trainers, resp.), and the absence or presence of the newly developed kinematic instrument-object interaction (i.e. the -I, and the -II setups, resp.). By this study design, the influence of these simulator features on the performance of participants could be compared.

Outcome Measures

The movements of the tip of the instruments were recorded during the pre- and post-test with the TrEndo tracking device, developed at Delft University of Technology[Chmarra et al., 2006], and motion-analysis parameters were established. The motion-analysis parameters were:

- » Time: defined as the total time taken to perform the task (s)
- » Total path length: defined as the average length of the curve described by the tip of the right and the left instrument while performing the task (m)
- » Motion in depth: defined as the total distance travelled by right and left instrument along its axis (m)

Time expresses the speed with which the exercise has successfully been performed. Path length is a measure for the economy of movements. The motion in depth is influenced by the depth perception of the trainee, in which problems with perceiving depth is likely to result in a longer motion in depth. Outcome measures were the differences between the parameters at the pre- and the post-test.

Statistical analysis

The recorded pre- and post-test results were collected, and analysed with the Statistical Package for Social Sciences (SPSS, version 16.0, Chicago, IL). The median and range of the outcome measures were given in case the data were not normally distributed. The relative improvement in parameters was calculated for the individual participant, and was expressed in percentage of the pre-test score. Additionally, the mean improvement within each group was determined. The Wilcoxon signed-rank test was used to establish the difference between pre- and post-test results. A p-value less than .05 was considered statistically significant.

RESULTS

In total 50 novices were enrolled in the study, and completed the entire study protocol. All denied prior laparoscopic or simulator experience. The five groups did not differ significantly with respect to gender, percentage of right-handed persons, self-perceived dexterity, and history of computer gaming.

The median scores and ranges of the pre- and post-test results are presented (Table 1). No statistically significant differences were present between the five groups regarding the pre-test results.

The observed improvement varied among groups. The control group did not show a significant improvement at post-testing with respect to time, path length and motion in depth. Regarding the four training modalities, all groups significantly improved in time to completion of the rubber band task. Regarding both economy of movement parameters, path length as well as motion in depth improved significantly in both box trainer groups. The VR-II trained group also improved significantly with regard to both these parameters, but the VR-I trained group did not.

Table 1. Time and economy of movement parameters.

	Pre-test median	(range)	Post-test median	(range)	Improvement p-value
VR1 (n=10)					
Time [s]	257	(240 - 345)	158	(117-277)	<.005
Path Length [m]	7.4	(4.8 - 22.3)	5.1	(4.0 – 17.2)	N.S.
Motion in Depth [m]	2.3	(1.1 - 4.1)	1.8	(1.2 – 4.6)	N.S.
VR2 (n=10)					
Time [s]	204	(152 - 413)	173	(130-311)	<.005
Path Length [m]	7.5	(3.5 – 14.8)	5.2	(3.3 – 10.9)	<.05
Motion in Depth [m]	2.2	(1.3 – 4.3)	1.8	(1.2 – 3.0)	<.05
Box 1 (n=10)					
Time [s]	245	(160-490)	156	(133-250)	<.005
Path Length [m]	8.1	(5.7 – 13.1)	4.9	(3.6 – 6.9)	<.005
Motion in Depth [m]	2.4.	(1.3 – 4.1)	1.7	(1.3 – 2.6)	<.01
Box 2 (n=10)					
Time [s]	245	(189-399)	157	(130-233)	<.005
Path Length [m]	6.8	(5.4 – 9.4)	4.8	(3.9 – 10.1)	<.05
Motion in Depth [m]	2.1	(1.6 – 3.1)	1.5	(1.0 – 2.5)	<.05
Control (n=10)					
Time [s]	255	(97-499)	195	(115-432)	N.S.
Path Length [m]	5.7	(3.0 – 1.5)	5.0	(3.4 – 15.6)	N.S.
Motion in Depth [m]	1.9	(1.2 – 3.8)	1.6	(0.9 – 4.1)	N.S.

Improvement is calculated using the Wilcoxon signed-rank test on the difference post-test and pre-test. VR-I: set-up in conventional VR environment, VR-II: set-up with kinematic object interaction application, Box-I: box trainer equivalent of VR-I, Box-II: box trainer equivalent of VR-II. Control: no training

DISCUSSION

Box training leads to a significant improvement in speed and in economy of movement during an exercise in which both force application and hand eye coordination are required. Conventional VR training results in improvement in terms of speed alone. However, a VR setup supplied with additional kinematic instrument-object interaction has an enhanced training capacity which is shown by the significant improvement in economy of movements of the trainees.

Prior studies have already compared the learning potential of box trainers and VR trainers.[Chmarra et al., 2008; Hamilton et al., 2002; Jordan et al., 2000; Kothari et al., 2002; Madan et al., 2007; Munz et al., 2004; Pearson et al., 2002; Torkington et al., 2001b] Most of these studies did not reveal significant differences in outcome measures.[Kothari et al., 2002; Madan & Frantzides, 2007; Munz et al., 2004; Pearson et al., 2002; Torkington et al., 2001b] Two studies showed an advantage for the VR trainer,[Hamilton et al., 2002; Jordan et al., 2000] and one showed an advantage for the box trainer.[Chmarra et al., 2008] An important limitation of the majority of these studies is that the tasks were not equivalent in the compared trainers. As the training conditions were unequal, the implication of these studies' results is limited. However, in the study of Chmarra et al., novices performed three equivalent exercises in both a VR trainer and a box trainer using a cross-over design.[Chmarra et al., 2008] They found that VR trained novices perform worse in a box trainer than the non-trained group who started with box training for the one exercise in which force transmission was required. For the exercises that mainly require hand eye coordination the group that had been trained on VR outperformed the non-trained group. These results indicated the effect of the need of realistic feedback to train tissue handling.

The novelty of our study is that not only equivalent exercises were used for all training setups, but also a pre- and post-test that differed from the trained task. Regarding the latter, individual progression can be taken as the result of the training setup, combined with a fixed effect of having performed the pre-test. Moreover, by choosing a force requiring task, we intended to simulate tissue handling during laparoscopy instead of only hand-eye coordination. However, the results of the box training groups might have been positively influenced by the fact that pre- and post-test are performed in a box trainer. By choosing equivalent exercises in this study, it was intended to have the presence of haptic feedback and kinematic instrument-object interaction as the only varying features.

Large ranges in pre-test scores with skewed distribution were observed. This can be explained by a variance in innate ability. Due to this distribution, it was only possible to draw conclusions about whether each training setup led to a significant improvement. Unfortunately, no quantitative comparison between the training systems could be made. Though not statistically proven, natural haptic feedback seems superior to a VR trainer with the newly developed interaction, as indicated by the larger percentage of improvement in economy of movement parameters in both box trainer setups when compared to the VR-II setup (median: 36 vs. 26% in path length, and 30 vs. 12% in motion in depth for both box trainers groups and VR-II, resp.).

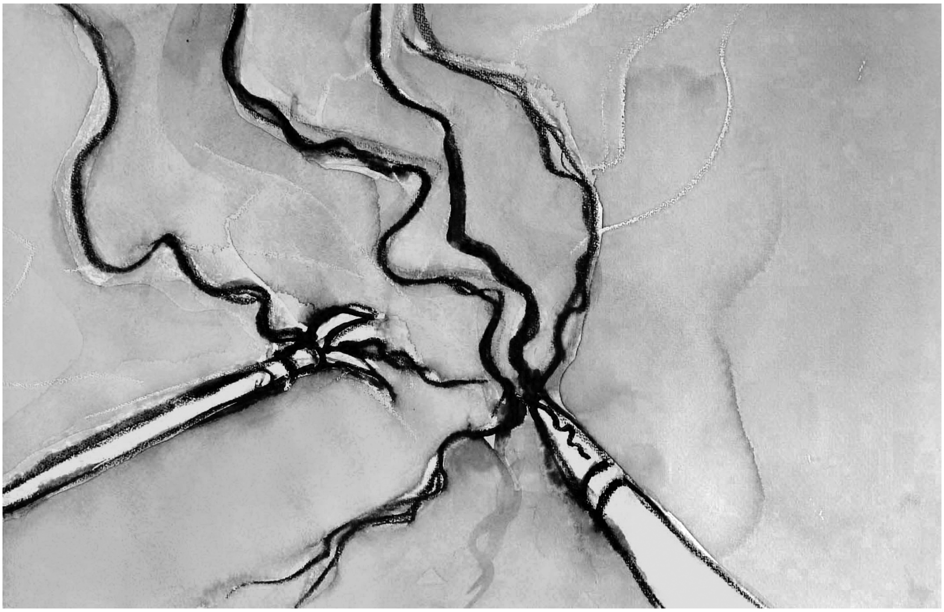
Continued training is required to achieve real competence in basic laparoscopic skills. However, it is found that much progress is generally made during the early phase of the process to acquire psychomotor skills.[Larsen et al., 2006] Therefore, despite the short duration of the training, a significant progression in psychomotor skills could be observed. An additional advantage of a short duration of the training is that the experiment could be held in one session without fatigue of a participant influencing the results.

Haptic feedback is considered to be essential for tissue handling. Next to providing information on tissue texture, shape and consistency, it can be used to regulate force application and to avoid tissue damage.[Strom et al., 2006] In laparoscopy, the balance between a firm grip on the tissue and not causing any damage even is harder to acquire.[Westebring-van der Putten EP et al., 2008] From this theoretical point of view, training using a model with haptic feedback should be considered superior in order to acquire proper force application. On the other hand, new technologies like robotic surgery are introduced in the clinical field. Probably, training models without haptic feedback will provide surgeons with good psychomotor skills to become proficient in this technique.

The transferability of skills acquired on simulators to the real OR setting remains the key concern, though the hardest to objectify. This transferability to laparoscopic surgery was proven for box trainers[Scott et al., 2000] as well as for VR trainers[Grantcharov et al., 2004; Seymour et al., 2002], using global rating scales and expert opinions. Based on our study on simulator features and on theoretical considerations, we judge a box trainer system with a natural instrument-tissue interface to be superior to VR training systems for acquiring tissue handling skills in laparoscopic surgery. Furthermore, box trainer are cheaper and easy accessible, which makes them likely to be actually used for laparoscopic skills training.[Sharma et al., 2009] However, if a VR training system is selected to train these skills, a system with kinematic instrument-object interaction can be a promising surrogate for haptic feedback to train tissue handling.

CHAPTER 4

OPTIMIZING LAPAROSCOPIC SKILLS TRAINING: DOES A FIXED CAMERA COMPROMISE DEPTH PERCEPTION?



Ellen Hiemstra
Navid Hossein pour Khaledian
John van den Dobbelsteen
Jenny Dankelman
Frank Willem Jansen

Submitted

INTRODUCTION

In laparoscopic surgery, the image display system is the only visual interface between the surgeon and the operation field. Inherently, the three-dimensional (3D) operation field has to be perceived from a two-dimensional (2D) screen. Additionally, by looking at the monitor, the surgeon indirectly observes his hands manipulating the laparoscopic instruments. This may result in perceptual disturbances and distorted hand eye coordination [Heemskerk et al., 2006].

The camera is a substitute for the surgeon's eyes, and therefore, its position and navigation are of utmost importance. Many studies have been conducted to reveal the influence of camera and the operative setup on laparoscopic performance and the surgeon's workload in simulator settings[Ames et al., 2006; Conrad et al., 2006; Emam et al., 2002; Hanna et al., 1998; Haveran et al., 2007; Matern et al., 2005; Moschos et al., 2004; Omar et al., 2005; Smith et al., 2005; Zehetner et al., 2006]. The monitor should to be positioned in front of the surgeon, preferably in a gaze-down position[Hanna et al., 1998; Haveran et al., 2007; Matern et al., 2005; Omar et al., 2005; Zehetner et al., 2006]. Furthermore, a 0 degree scope results in the best performance during laparoscopy, and even the modest alteration in perspective results in a deterioration of performance[Ames et al., 2006; Omar et al., 2005]. The rotational angle of the laparoscopic image to the true horizon must be kept to a minimum to maintain a stable horizon, and for an optimal performance [Conrad et al., 2006]. Finally, it was found that the best place for a surgeon to stand is right in front of the laparoscopic instruments [Moschos & Coleman, 2004].

During the experiments described above, the camera was always placed in a fixed position. In laparoscopic practice, however, the camera is often navigated by either the surgeon who performs one-handed surgery, or by an assistant under direct oral instruction of the surgeon who performs two-handed surgery. Instable camera movements result in fatigue of the surgeon and delays in operative times [Bennett et al., 2011; Heemskerk et al., 2006], but in general, camera navigation provides the surgeon with 'depth cues'. For example, objects closer to the camera "move" faster as a result of navigation, than objects further away. The subsequent better understanding of the 3D operation field will facilitate hand-eye coordination and thereby efficient instrument-movements.

Box trainers are designed for basic laparoscopic skills training. However, many box trainers are supplied with fixed camera systems, despite the theoretical importance and the practical application of a navigated camera. To our knowledge, no research has been conducted on whether a camera navigation setup influences the proficiency gaining process. Therefore, this study compares a fixed camera to a navigated camera during laparoscopic skills training in this study. We try to answer the question whether a fixed camera position compromises depth perception during laparoscopic skills training.

MATERIAL AND METHODS

The study was conducted at the skills laboratory of the Leiden University Medical Center (LUMC) in the Netherlands. An inanimate box trainer was used with a separate monitor and an endoscope with 0 degree camera In this box trainer, a validated beads placing task [Kolkman et al., 2008] had to be performed (Figure 1). To fulfil this task, the beads had to be placed in a designated position using a laparoscopic grasper with the (dominant) right hand. For this exercise a correct interpretation of the 3D operation field is indispensable.

Participants

Right-handed medical students, in the preclinical phase of their study, and without prior experience with laparoscopic surgery or training (novices) were recruited to the study. Participation was on a voluntary basis. After enrolment, all participants completed a questionnaire providing demographic information (i.e. gender, self-perceived dexterity, and computer gaming experience) in order to compare baseline characteristics.

Study design (Figure 2)

Each novice was randomly assigned to eight beads placing tasks in one of the following three camera navigation setups, thereby testing the influence of camera navigation on laparoscopic skills acquisition:

- I: Assistant navigated camera: the task was performed with the dominant right hand while the camera was navigated by an assistant who was positioned on the left side of the participant.
- II: Self navigated camera: the task was performed with the dominant right hand while the camera was navigated by the participant's left hand.
- III: Fixed camera: the task was performed with the right hand while the camera was fixed in a standardized position.

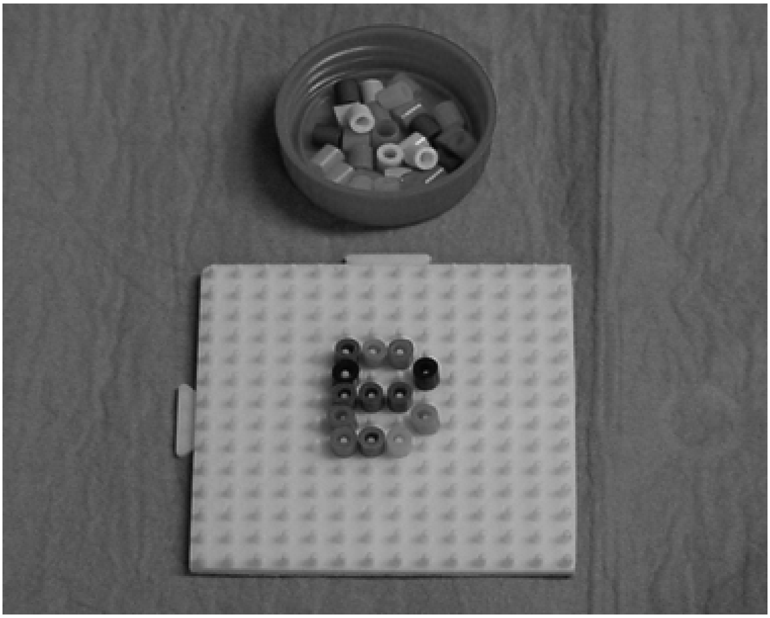


Figure 1. Beads placing task.

Randomization was done by using the website www.randomization.com, with an equal distribution of the participants to the three groups. Regarding setup III, camera navigation was performed only on demand, and always by the same assistant (NH). Possible commands were: *centre the work field, zoom in,* and *zoom out*. Instructions were given in advance.

Each participant performed the beads placing task eight times in order to gain insight in the learning curves over time with the different camera setups, instead of only comparing the performance at the start of the skills training. The eight trials were distributed over two sessions of four trials with approximately one week in between. This distributed training was chosen in order to prevent a worse performance due to fatigue.

Outcome measures

The movements of the tip of the laparoscopic grasper, used for picking up and transporting the beads, were tracked using a built-in tracking system, the TrEndo. The TrEndo, developed at the Delft University of Technology, allows realistic movements of the laparoscopic instrument in four degrees of freedom and real-time recording of the instruments movements [Chmarra et al., 2006]. Time (seconds) to a successful completion of the task was recorded and used as outcome measure. Additionally, two kinematic parameters were calculated using the recorded movements: the total path length (meters), and the motion in depth (meters). Total path length was defined as the total distance the tip of the instrument travelled, and motion in depth is defined as the total distance travelled by the instrument along its axis. The latter parameter was chosen as it might be indicative for a trainee's depth perception[Cotin et al., 2002].

Statistical analysis

Data were recorded and analysed in SPSS 16.0 software package (SPSS, Chicago, IL, USA). The time and the motion in depth were plotted for each trial. For comparison of baseline characteristics of the three groups, a Student t-test was used for normally distributed continuous variables, and a Pearson's Chi-square to test dichotomous data. A mixed design ANOVA was used in order to compare the effect of camera navigation setup and skills training. The model contained a between-subject factor for the difference in camera setup and a within-

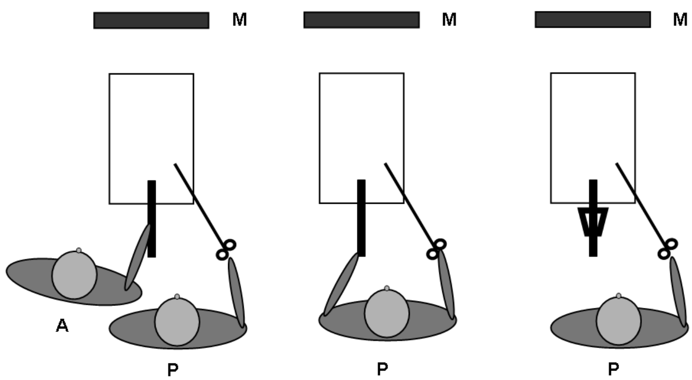


Figure 2. Study Design. I: Assistant navigated camera, II: Self navigated camera III: Fixed camera. (A = assistant, P = person, M = monitor).

subject factor for the repeated measurements (i.e. trials) for each participant. We determined the effects of these independent variables on the outcome measures time, path length and motion in depth. Bonferroni post-hoc tests were performed to determine whether there were significant differences between the trials. A p-value less than .05 was considered statistically significant, 95 per cent confidence intervals (95%CI) were calculated.

RESULTS

In total, 69 right-handed novices were enrolled in the study. None of them had prior surgical experience. They were equally distributed among the camera setups: a self-navigated camera (n=23), a researcher-navigated camera (n=23) and a fixed camera (n=23). They all completed the entire study protocol of eight trials. Among the participants, 21 were male and 48 were female, 15 among them did frequently play video games. No variance with respect to these parameters was observed among the three groups. Also with respect to self-perceived dexterity, participants had been equally distributed among the three groups.

The box plots for each of the three camera setup groups are displayed. (Figure 3) Time, path length and motion in depth improved for all three groups of participants within the eight trials ($p<.001$ for all three parameters). Post-hoc testing showed that the performance significantly improved in the first three trials but that performance was about equal for the following trials.

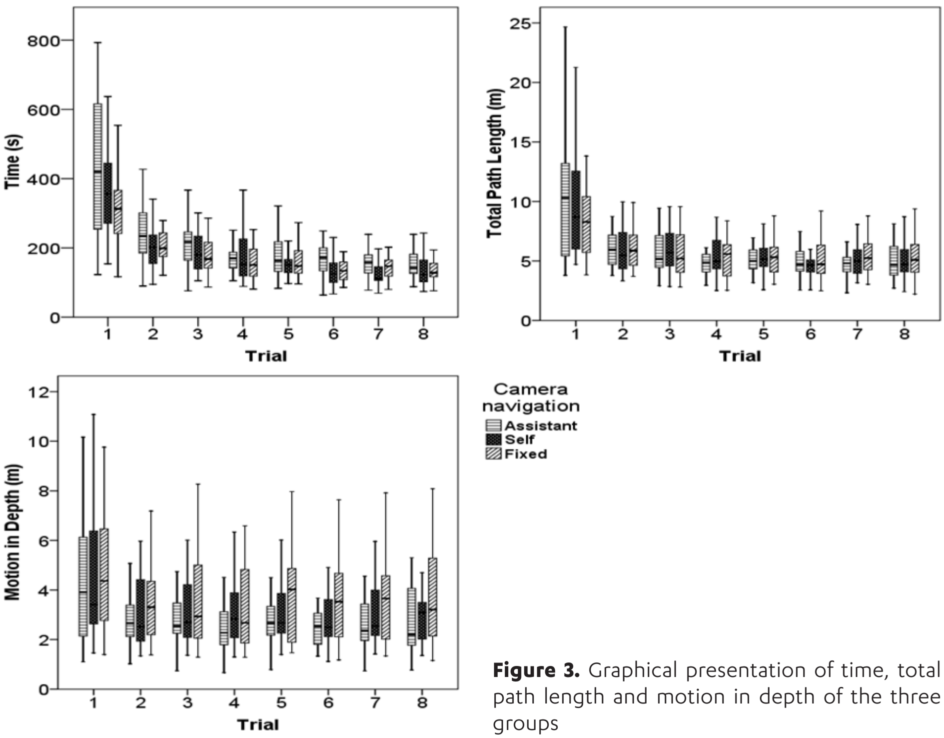


Figure 3. Graphical presentation of time, total path length and motion in depth of the three groups

The camera setup was not a significant factor. Despite this lack of difference of camera setup on performance, the group with the assistant navigated camera tended to need more time for completion of the task during the first trial than the group with the self-navigated and the fixed camera position. Additionally, the motion in depth tended to be longer for the fixed camera group during the last four trials than for the other two setups. In other words, the fixed camera group tended to travel a longer path along the instruments axis.

DISCUSSION

The camera setup does not influence the performance of a task requiring hand-eye coordination, with respect to time, path length or motion in depth. As a result, the hypothesis is not confirmed that a fixed camera position compromises depth perception.

However, motion in depth tends to be longer in the fixed camera group compared to the self- and the assistant navigated camera group during the last four trials. Therefore, it is possible that the camera setup plays a role during basic laparoscopic skills training, although this was not apparent from our study. It has to be considered that the beads placing task is relatively simple. Repetitive movements from the basket to the pegboard have to be carried out. Possibly, this has helped all groups to correctly interpret the 3D operation field. On the other hand, it is a validated task in which proper hand-eye coordination is of utmost importance, because every bead should be placed in the exact indicated position, without causing the beads already placed to fall over. As a consequence, this task seemed appropriate for answering our research question. Maybe, an influence of camera navigation setup will be revealed during more complex tasks, like intracorporeal knot tying. That is in accordance with the findings of Omar et al. who proved that a better monitor stance leads to more improvement in performance of more complex tasks [Omar et al., 2005].

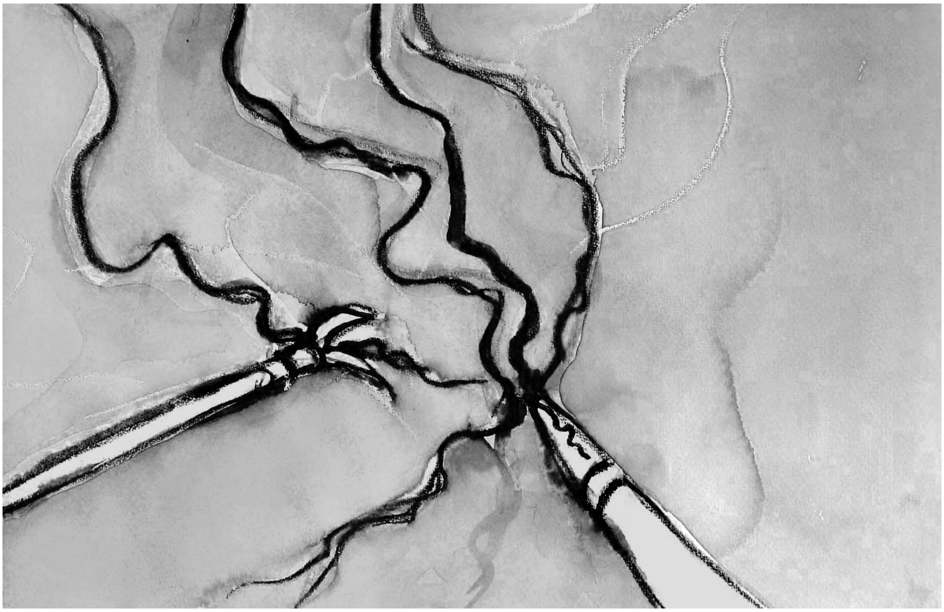
Inherently to the absence of difference between the three groups, no difference was observed between camera navigation by the trainee and by an assistant. However, a tendency was present that the participants whose camera was navigated by the assistant required more time during the first trial. A theoretical explanation is that the participants needed time to instruct the assistant, combined with getting familiar with the exercise.

In laparoscopy, a surgeon needs to be proficient in camera navigation and in giving clear instructions to the assistant who navigates the camera. In both ways he should be able to interpret the 3D operation field. As a consequence, it is worthwhile to train camera navigation skills in a simulator setting. A box trainer model and a virtual reality trainer both have been validated and found to be effective training tools for that purpose [Bennett et al., 2011; Korndorffer, Jr. et al., 2005b]. Despite the importance of being able to perform surgery with a navigated camera system, also standards are being used in order to fix the camera position. The major advantage of a standard is the stability of the view. The decision for a system will depend on the surgeon's preference and the required changes of the operation field during a procedure. For example, during a long lasting laparoscopic nephrectomy the required stability of the image may outweigh the need to change the camera position frequently, whereas it is the other way around during the laparoscopic removal of an adnexal cyst.

In conclusion, the influence of camera navigation is either absent or too small to be observed during a basic laparoscopic skills task. A fixed camera setup allows solitary training without clearly compromising depth perception, while training with a navigated camera will lead to enhanced navigation skills, and to training of communication skills with an assistant. Therefore, both ways of training have their own benefits and from a clinical point of view the combination seems superior.

CHAPTER 5

INTRACORPOREAL SUTURING: ECONOMY OF MOVEMENTS IN A BOX TRAINER MODEL



Ellen Hiemstra
Magdalena K. Chmarra
Jenny Dankelman
Frank Willem Jansen

Adapted from JMIG, 2011; 18: 494-9

INTRODUCTION

Training outside the operating room (OR) prior to patient exposure is important to progress along the first part of the learning curve in order to enhance patient safety. Preclinical practice using inanimate models improves psychomotor skills and translates into improved performance in the OR.[Grantcharov et al., 2004; Scott et al., 2000; Seymour et al., 2002] Among the laparoscopic exercises that can be learned using a simulator, intracorporeal suturing is unique, because it is directly applicable in clinical practice. Moreover, all basic laparoscopic skills are incorporated in this task, i.e. ambidexterity, judging depths, handling materials, manipulating instruments, and using fluid movements.[Rosser et al., 1997] Additionally, being proficient in suturing, inclusively being able to tie an intracorporeal knot, is a prerequisite to perform advanced laparoscopic surgical procedures.[Aggarwal et al., 2006] Suturing skills are needed if a complication occurs (e.g. a bleeding, a lesion in the urinary bladder or the intestine), or in case of a dysfunction of suturing devices.

Intracorporeal suturing can be practiced on two categories of simulators: computerized virtual reality (VR) trainers and physical box trainers, which are also called video trainers. The latter trainers have been criticized for the lack of objective assessment of movements, and for the low fidelity of most exercises with regard to real laparoscopic procedures.[Aggarwal et al., 2004] However, Grober et al. revealed that low-fidelity models can be as efficient as high-fidelity training models for technical skills acquisition.[Grober et al., 2004b] In fact, the advantage of box trainers over VR trainers is that real laparoscopic instruments, camera and monitor can be used, which result in natural haptic feedback and perceptions of depths. Regarding suturing, the use of a needle holder and the various suturing threads and needles can be practiced. Furthermore, the absence of instant feedback on economy of instrument movements in box trainers recently has been overcome by the development of tracking devices.[Chmarra et al., 2006] By consequence, objective assessment of economy of movements is facilitated during intracorporeal suturing in box trainers.

Objectifying a subject's level of performance during laparoscopic tasks is worthwhile for skills acquisition to allow a continuous refinement, as well as for credentialing or certification purposes. However, prior to the implementation of a tracking device for assessment purposes for a specific task, the construct validity (i.e. the ability to discriminate between clinicians of a different skills level) of the economy of movement parameters needs to be established, and benchmark criteria need to be set. Two studies have been performed to validate economy of movement parameters during a suturing task in a box trainer.[Aggarwal et al., 2006; Van Sickle et al., 2005] Aggarwal et al. used a tracking device that needs to be applied on the dorsum of a surgeons hand to obtain dexterity data[Aggarwal et al., 2006], and is relatively time consuming. Van Sickle et al. studied a needle driving task, but did not focus on the knot tying.[Van Sickle et al., 2005] A suture, however, does not function without a proper knot.

The current study is conducted to establish the construct validity of time and three economy of instrument movement parameters for the entire suturing task recorded by an easy applicable tracking device, the TrEndo.[Chmarra et al., 2006] Additionally, the improvement of the movement parameters is compared to the improvement in time to complete the task during the three trials. Subsequently, an expert standard is set.

MATERIALS AND METHODS

Measurements were performed in the skills laboratory of the Leiden University Medical Center (LUMC) in the Netherlands.

Participants

For this study, novices, intermediates and experts were recruited. The novices were medical students in the preclinical phase of their study, and consequently without prior operative experience. They were recruited by means of advertisement in the medical library of the LUMC. Intermediates were residents in Obstetrics and Gynaecology in all 6 post-graduate years (PGYs). For their recruitment, an email was sent to all residents who attended the Obstetrics and Gynaecology specialty training at the LUMC. All are being trained in performing laparoscopic surgery, and are two-monthly scheduled to train their basic laparoscopic skills in a laboratory setting. Experts were experienced surgeons in minimally invasive surgery (MIS) who met the following three preconditions:

- Their experience exceeded 200 laparoscopic procedures.
- Their surgical palette contained advanced procedures. Herein, the advanced gynaecological procedures are defined in the ESGE standard of laparoscopy.[ESGE 2009] For general surgery, we choose to set herniorrhaphy, fundoplication, colectomy, adrenalectomy, and splenectomy as examples of advanced procedures.
- Intracorporeal suturing was practiced in the clinical situation.

They were recruited at the department of surgery and of gynaecology at the LUMC, and during a nationwide assembly among gynaecologists who are involved in minimally invasive surgery.

Box trainer

The trainer consisted of a box, measuring 45 x 30 x 25 cm, with a non-transparent cover, and was designed and fabricated at the LUMC.[Kolkman et al., 2008] The image of a 0 degree scope was displayed on a monitor.

Task

The task involved the placement of a simple suture followed by tying an intracorporeal knot. The exercise started with the needle positioned in the needle holder. A thread length of 12 cm of 2/0 Vicryl was used, Ethicon, Johnson and Johnson. A proper bite had to be taken of the suturing pad in a pre-marked area. A 3-throw square knot had to be tied, after the needle and a substantial part of the suture material had been driven through. The knot tying technique was standardized. The suture had to be wrapped twice around the left hand needle holder, and the short end of the thread had to be pulled through the loops. Next, the dread had to be wrapped once around the right hand needle holder and the short end had to be pulled through this loop. Finally, the dread had to be wrapped around the left hand needle holder again and the short end of the thread had to be pulled through this final loop (Figure 1). Prior to performing the task, a demonstration video was shown, followed by the step-by-step graphical explanation (Figure 1). Next, the video was demonstrated a second time. All participants had to perform the suturing task three consecutively times. No practicing was permitted before, or in between, the suturing trials. If necessary, the researcher (EH) coached a participant to perform the correct next step, in order to keep every trial in the study, and to ensure that every knot was of good quality (no slippage was allowed) and was performed in a standardized way.

Motion analysis, four parameters

The movements of the laparoscopic instruments were recorded with the TrEndo tracking device in four degrees of freedom (DOFs): an up-down translation (1st DOF), a forward-backward (2nd DOF), and a left-right (3rd DOF) rotation around the incision point, and the rotation of the instruments around its longitudinal axis (4th DOF).[Chmarra et al., 2006] The recorded data were the **time** (s), defined as the total time taken to perform the task, and additionally the following three economy of movement parameters[Chmarra et al., 2006]:

- **Path length** (m): defined as the average path length of the right and the left instrument tip during the task;
- **Motion in depth** (m): defined as the average of the distance travelled by right and left instrument along its axis;
- **Motion smoothness** (m/s^3): a motion analysis parameter based on the third time derivate of position, which represents the change in acceleration. The motion smoothness was calculated by averaging this parameter for the right and the left instrument. A low-pass Butterworth filter was used to filter the raw data.

Hypothesis

The four motion analysis parameters (i.e. time, path length, motion in depth and motion smoothness) are able to distinguish surgeons of a different skills level. Especially, the parameters can discriminate between novices and experts during an intracorporeal suturing task.

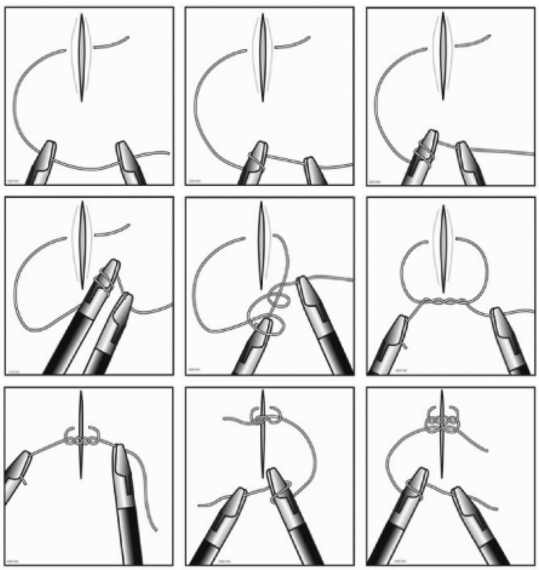


Figure 1. Standardized way to tie an intracorporeal knot.

Statistics

The data were analysed using the SPSS 16.0 software package (SPSS, Chicago, IL, USA). In order to test whether the motion-analysis parameters can discriminate between novices, intermediates and experts, non-parametric Kruskal-Wallis tests were used for each parameter. In case the groups differed significantly, pair-wise comparisons were performed to test the difference between each group using Mann Whitney. Additionally, the Bonferroni correction for multiple testing was used with respect to the differences between novices and experts, and therefore to test the primary research hypothesis. Finally, to test whether improvement within each group was present a Wilcoxon's signed rank-test was used, comparing the first and the third trial. Probability below .05 was considered statistically significant.

RESULTS

In total, 19 novices, 12 intermediates, and 11 expert laparoscopic surgeons participated in the study. All 19 novices who had responded to the advertisement denied previous surgical experience and completed the entire study protocol of 3 trials. Of the 22 residents who received an e-mail, 12 agreed to participate in the study: four in PGY1, two in PGY2, three in PGY3, two in PGY 4, and one in PGY 5. All completed the entire study protocol. Of the 11 participating laparoscopic experts, four were consultants in surgery, and seven were consultants in gynaecology. All met the inclusion criteria of the study definition of an ''expert.'' Eight experts completed the entire protocol of 3 consecutive suturing tasks. Two experts (one surgeon and one gynaecologist) did not perform the third trial, and one expert surgeon performed only the first trial. Time constraints were given as a reason to not complete the entire study protocol.

Examples of the typical trajectory of the right needle holder tip as used by a novice and an expert, both right-handed, are presented graphically in figure 2.

Performance of novices, intermediates, and experts throughout the three trials are displayed graphically in Figure 3 for the four motion-analysis parameters. Differences were observed between the three groups of participants ($p < .001$ during all trials for all four parameters). A Bonferroni correction for multiple testing revealed that performance of novices significantly differed from that of experts in all trials ($p < .01$ for time, path length, motion in depth, and motion smoothness). The difference between each combination of groups, calculated using the Mann-Whitney test, are shown for all parameters during the four trials (Figure 3).

The lower and upper bounds of the boxes represent the 25th and 75th percentiles, respectively, and the black line in the boxes indicate the median. *$p < .05$, **$p < .01$, and ***$p < .001$, all Mann-Whitney test.

Performance of novices and intermediates improved significantly for all four parameters (Table 1). However, no significant improvement was observed in performance of the experts. For novices, initial time improved, on average, by 42%, and path length, motion in depth, and motion smoothness improved by 26%, 25%, and 8%, respectively, of the initial score.

To set an expert standard for the 3-throw knot, the median scores of the 4 motion-analysis parameters of the experts during the second trial were used (Table 2).

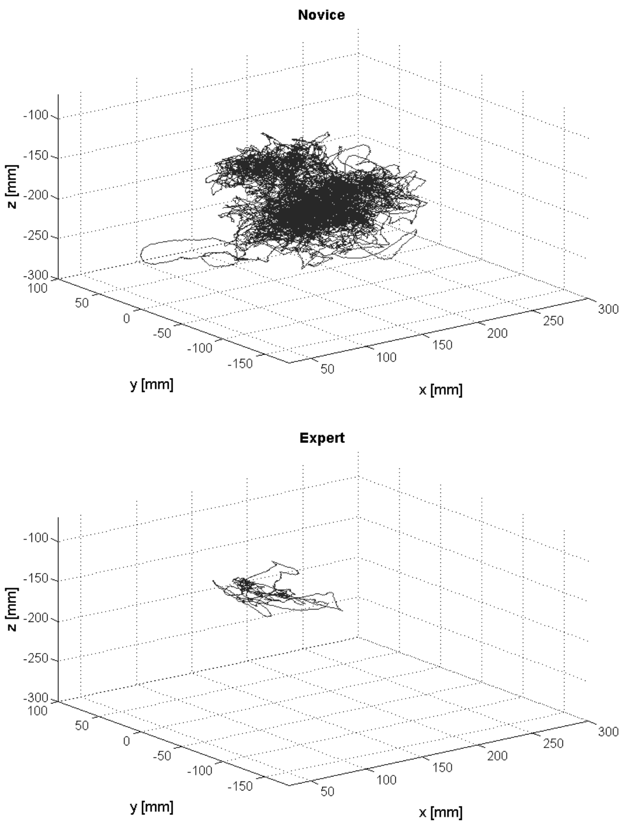


Figure 2. Typical trajectory (in millimetres) of the right-hand instrument during intracorporeal suturing by a novice (A) and an expert (B).

Table 1. Improvement in 4 parameters during 3 trials.

	Parameter	p Value trial 1 - 3
Novice	Time	<001
	Path Length	<01
	Motion in Depth	<01
	Smoothness	<01
Intermediate	Time	<01
	Path Length	<01
	Motion in Depth	<01
	Smoothness	<01
Expert	Time	NS
	Path Length	NS
	Motion in Depth	NS
	Smoothness	NS

Wilcoxon signed rank test was used to compare first and third trials.

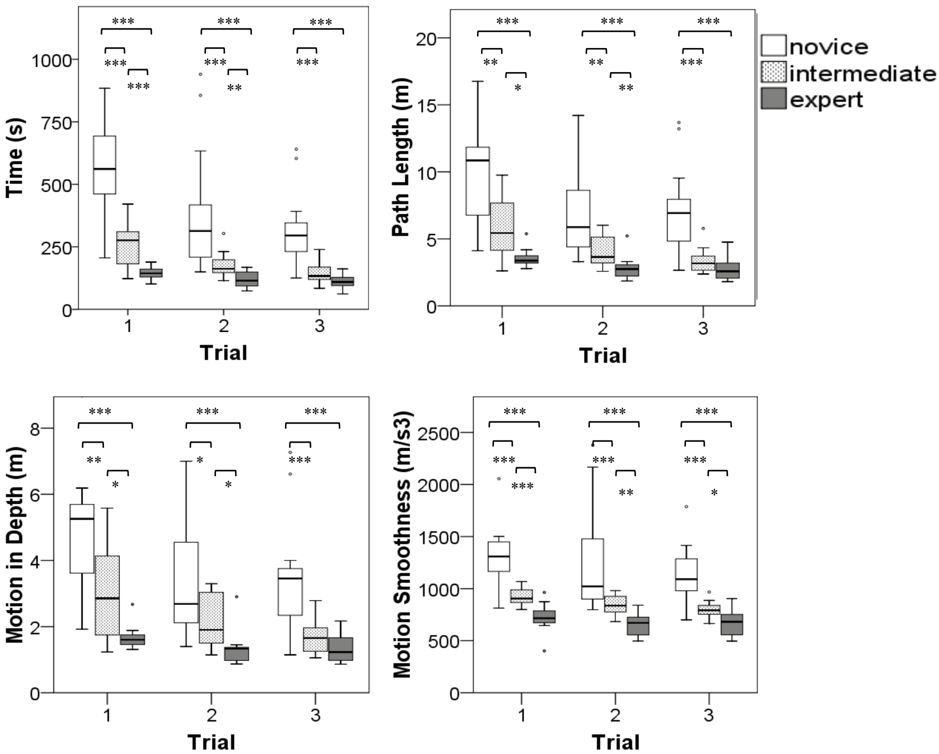


Figure 3. suturing task scores for novices, intermediates and experts.

Table 2. Expert standard, established during second trial.

Motion analysis Parameter	Median	Interquartile Range
Time [s]	114	90 – 150
Path Length [m]	2.75	2.17 – 3.13
Motion in Depth [m]	1.33	0.96 – 1.39
Motion Smoothness [m/s^3]	670	550 – 728

DISCUSSION

This study has confirmed the hypothesis to be true that the four motion-analysis parameters are able to discriminate between groups with different levels of experience during an intracorporeal suturing task in a physical box trainer. This indicates the construct validity of these objective assessment parameters for psychomotor skills for intracorporeal suturing.

Although experts outperformed intermediates during the first and second trial for all parameters, no significant difference was present during the third trial for time, path length and motion in depth. The intermediates were residents in Obstetrics and Gynaecology who had been exposed to one or more laparoscopic skills training sessions during their specialty training. As a result of this training, some of them may approach the expert level after a short rehearsal of two intracorporeal sutures. This might be an essential difference between being proficient (i.e. an expert), and progressing along a proficiency gaining curve (i.e. a resident).

The relevance of economy of movement parameters in the in vivo situation was revealed by their ability to discriminate between surgeons of a different level of experience during a laparoscopic cholecystectomy on a porcine model.[Smith et al., 2002] Furthermore, specified feedback on performance can be provided if these parameters are used during simulator training. Unfortunately, most of the currently described tracking devices can only be used in a virtual environment.[Chmarra et al., 2007] Regarding intracorporeal suturing, a recent study revealed no additional value of VR simulation over a box trainer model.[Botden et al., 2008] Unlike box trainers, VR trainers do not offer a natural instrument-tissue interaction, or the possibility to use real laparoscopic instruments and materials (e.g. a needle holder and suturing material). Determining motion-analysis parameters with the TrEndo in a box trainer combines the haptic advantages of box trainers with the objective assessment ability of VR trainers.

Prior studies have used time to completion of a suturing task to establish learning curves.[Kolkman et al., 2008; Munz et al., 2007; Van Sickle et al., 2005; Vossen et al., 1997] However, Smith et al. found that time improved during the first three trials and then stabilized, while the path length continued to improve over 10 trials.[Smith et al., 2002] Therefore, it was concluded that the learning curve for time alone fails to account for the more protracted learning curve for accuracy. We found that the novices' performance in suturing improved significantly for all four motion-analysis parameters across the three trials. However, the observed improvement in time was relatively larger (42%) than the improvement in the three economy of movement parameters (path length, motion in depth and motion smoothness). Since most of the training effect is achieved after tying 20-30 knots[Vossen et al., 1997], achievement of proficiency should not be expected within 3 trials. Additionally, it was revealed that experts still outperform novices after 100 trials to tie a proper knot.[Vossen et al., 1997] Our study is in line with that finding. The relatively small improvement in the economy of movement parameters, compared to the time to completion of the task, suggests that it takes more trials to perform an intracorporeal suture with efficient use of movements than to perform it quickly.

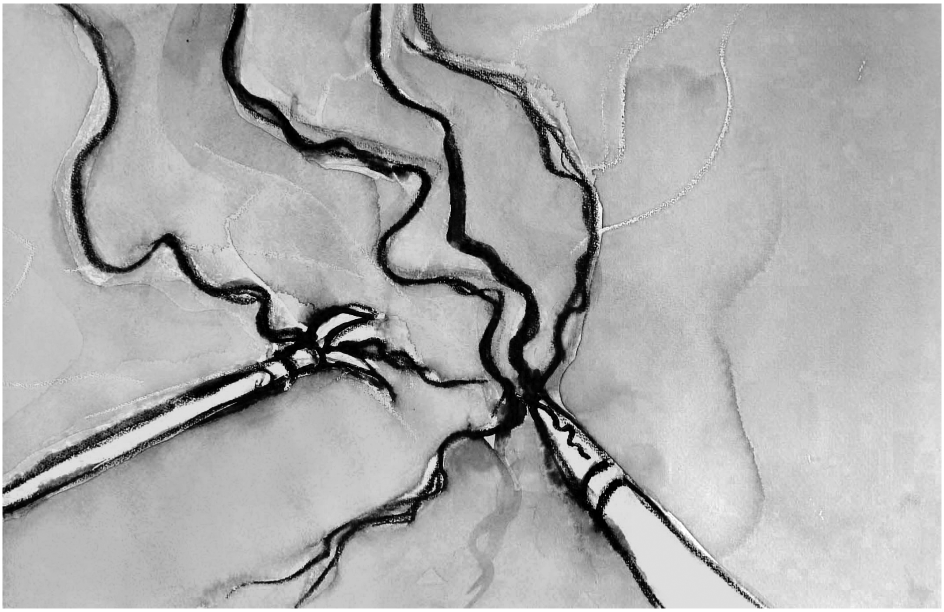
Another objective of this study was to set a performance standard for laparoscopic suturing by using the parameters of the experts' performance. Obviously, the experts did not improve significantly across the trials, since they had already achieved proficiency. Although arbitrary, we choose to set the experts performance at the second trial as the expert standard. The first trial was not chosen in order to correct for possible adaptation to the simulator setup, and the third trial had the disadvantage that it was only performed by 8 out of 11 experts. This performance standard can be used for training purposes, but also for assessment or even certification in order to enhance patient safety. Even though, experts are quite consistent in performance, expressed by small interquartile ranges of their scores, their spread around the median should be taken into account when this standard is implemented.

Intracorporeal suturing incorporates all basic laparoscopic skills[Rosser et al., 1997], and it is a prerequisite for advanced laparoscopic procedures as it is needed to handle possible complications or in case of instrument failure. Acquisition of suturing translates directly to the clinical situation. Concerns have risen whether the suturing skill is too complex to acquire for junior residents. However, the results of our study confirms those of Aggarwal et al. who revealed that residents with little or no previous laparoscopic experience are able to carry out the task competently after a short training course.[Aggarwal et al., 2006] Consequently, in our opinion, intracorporeal suturing training should be incorporated early in the residency curriculum.

The construct validity of time, and three economy of movement parameters (i.e. path length, motion in depth, and motion smoothness) has strongly been suggested for the intracorporeal suturing task performed in a box trainer. An expert standard has been set. Economy of movement parameters should be added to assessment package in order to form a framework of specified feedback, and to allow continuous refinement of psychomotor skills during training, and can be used for assessment and for certification purposes.

CHAPTER 6

RETENTION OF BASIC LAPAROSCOPIC SKILLS AFTER A STRUCTURED TRAINING PROGRAM



Ellen Hiemstra
Wendela Kolkman
Mathijs .A.J. van de Put
Frank Willem Jansen

Adapted from Gynecol Surg 2009; 6 229-35

INTRODUCTION

Laparoscopic surgery requires skills that are different from those required for open surgery. Simulators were developed to train these skills in a pressure-free environment with or without supervision [Munz et al., 2004]. They can roughly be divided into box trainers (also video trainers) and virtual reality trainers. The interest in training facilities outside the operating room (OR) was further enhanced by issues like quality control, patient safety, and cost-effectiveness[Feldman et al., 2004b; Munz et al., 2004]. Simulator training is shown to be effective in providing skills that are transferable to the OR[Anastakis et al., 1999; Fried et al., 1999; Schijven et al., 2005; Seymour et al., 2002; Torkington et al., 2001b] and to decrease procedural complications [Cadeddu et al., 2001; Martin et al., 1998]. Besides, simulators have potential for objectively assessing laparoscopic skills [Fried et al., 2004].

In the Leiden University Medical Center (LUMC) we developed a skills laboratory with an inanimate five-task laparoscopic simulation model (box trainer) for basic training and evaluation. In a recent study, our group has established the construct validity of these five tasks (i.e. the ability to discriminate between different skills levels)[Kolkman et al., 2008]. Additionally, the median score of five laparoscopic experts was set as performance standard for training and evaluation purposes[Kolkman et al., 2008]. This standard was based on one trial for all exercises. Finally, it was found that novices' skills improved quantifiably and met the performance standard within seven trials (after a 5-weeks training course)[Kolkman et al., 2008].

In order to establish an efficient laparoscopic training program, the retention of the acquired skills is highly important. Additionally, insight into retention is important to judge the necessity and frequency of continued training needed to maintain the acquired skills level.

As yet, a couple of studies have described skill retention after simulator based laparoscopic training [Grober et al., 2004a; Sinha et al., 2008; Stefanidis et al., 2005; Stefanidis et al., 2006b; Torkington et al., 2001a]. The retention test results varied considerably, as well as the study designs. On one hand, a 25% skills deterioration was observed three months after a one-day hands-on training course on a box trainer[Torkington et al., 2001a], and on the other hand excellent skills retention was revealed six months after validated proficiency-based training sessions[Stefanidis et al., 2006b].

The objective of the current study is to investigate the retention of skills one year after the start of our laparoscopic training program and to enhance the insight into the retention process.

MATERIAL AND METHODS

The study was performed in the skills laboratory located in the Department of Gynaecology at the LUMC in The Netherlands. The simulator was designed (FWJ) and fabricated at this tertiary teaching hospital. It consisted of an inanimate five-task box trainer with a non-transparent cover, measuring 45 x 30 x 25 cm using a 0º scope.

Outcome measures

The individual's performance on the box trainer was measured using a scoring system that rewarded precision and speed. During each task, the time to completion (seconds) and penalty points were measured. Scores were calculated by the addition of completion time and penalty points, thus rewarding both speed and precision (score = time + penalty points). Consequently, faster and more accurate performance was rewarded with a lower score. Additional to separately scoring each task, a sum score was calculated (the sum of scores of all five tasks).

Tasks

The tasks in this study, as well as the scoring system, were based on the studies of Derossis et al.[Derossis et al., 1998] and are shown (Figure 1). The tasks vary from simple placing object tasks to more complicated manoeuvres such as cutting and knot tying.

6

Pipe cleaner

This task involved the placement of a pipe cleaner though four small rings. A penalty was calculated when a ring was missed. Score = time in seconds + (the number of missed rings x 10).

Placing rubber band

This task required the participant to stretch a rubber band around 16 nails on a wooden board. A penalty was calculated when the rubber band was not stretched around a nail at the end of the task. Score = time in seconds + (the number of missed nails x 10).

Placing beads

This task involved the individual's placing 13 beads to form a letter 'B'. A penalty was calculated when a bead was dropped next to the pegboard. Score = time in seconds + (the number of dropped beads x 10).

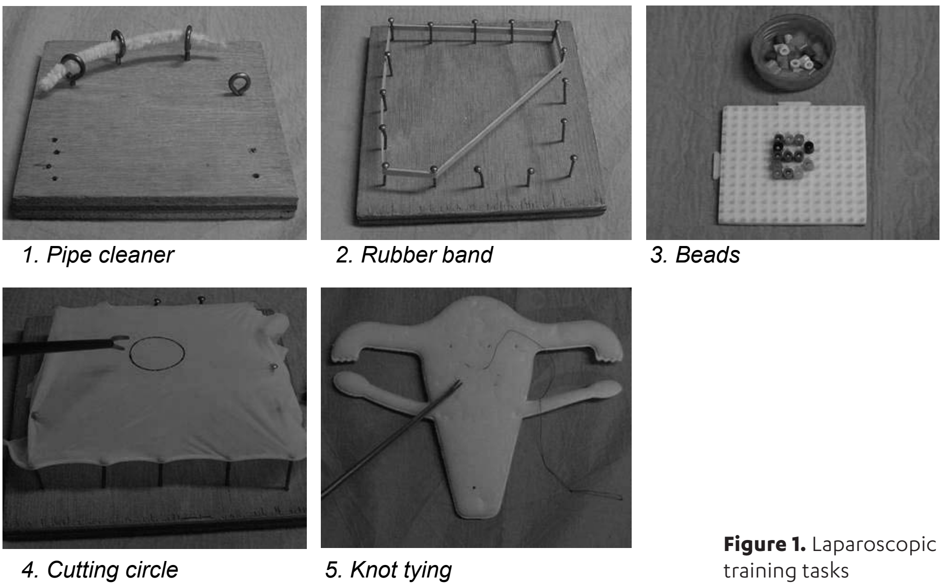


Figure 1. Laparoscopic training tasks

Cutting circle

This task required the participant to cut a circle from a rubber glove stretched over 16 nails in a wooden board. Penalty points were calculated when the individual deviated from cutting on the line. Score = time in seconds + surface of glove in milligrams deviated from circle.

Intra-corporeal knot tying

This task involved the tying of an intra-corporeal knot (two turn, square knots) in a foam uterus. A penalty was calculated to reflect the security (slipping or too loose) of the knot. Score = time in seconds + 10 when knot was slipping or loose.

Participants and measurements

The same eight medical students (novices) who had volunteered to participate and had been trained in our previous study[Kolkman et al., 2008] were asked to participate in the current study for a retention of skills test one year after the start of the training program. At the time of training they were in the second to fourth years of their medical study at the LUMC and had no prior experience with simulator training or clinical laparoscopy. A precondition for current participation was that they had not further been training or practicing their laparoscopic skills during the consecutive one year period.

As described previously[Kolkman et al., 2008], the novices had underwent baseline testing on the simulator followed by five weekly training sessions and had been measured again the week afterwards for final testing, as shown in figure 2. Novices performed all five tasks once during baseline testing, the training sessions and the final testing. Consequently, the novices had completed all tasks a total of seven times by the end of the study (one baseline test, five

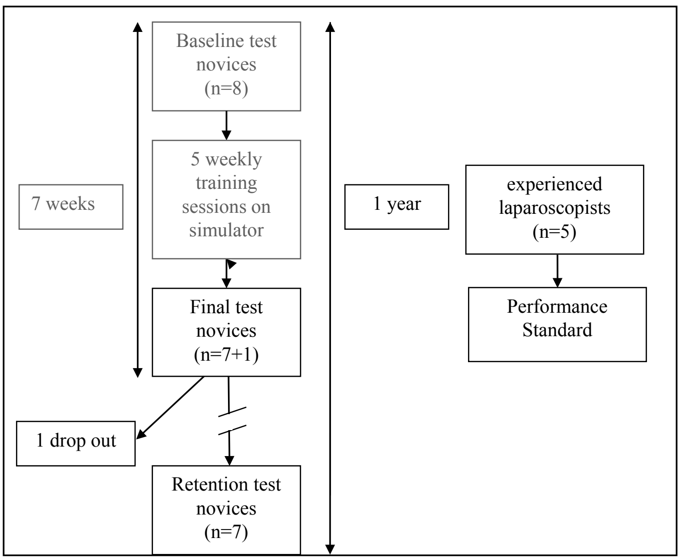


Figure 2. Flow chart of the study.

training sessions, one final test). Therefore, it will be referred to as seven trials. One year after the start of the training, they were asked to volunteer to perform the five tasks once more (retention test) for current study.

The primary outcome measure for the durability of the acquired skills was the comparison of the novices' retention test scores to their final test scores. Secondarily, the retention test scores were compared with the baseline test, and retention test scores were also compared with the performance standard. This standard was established by the median scores of five "expert" gynaecologists (having performed more than 100 advanced laparoscopic procedures) obtained from one single performance on the five simulator tasks once[Kolkman et al., 2008].

Statistical analyses

Collected data were analysed by SPSS 16.0 software package (SPSS, Chicago, IL, USA). Statistical analyses were performed using Mann-Whitney and Wilcoxon signed-rank test. Probability below .05 was considered statistically significant.

RESULTS

Among the eight students who participated in our initial study, one was unable to participate in the current study due to absence during the retention test. As a result, the performances of seven students who participated in our initial seven-week training program were measured on the box trainer for assessment of retention of basic laparoscopic skills. The participants were considered novices since they had no surgical, laparoscopic or simulator experience prior to the training program. Novices' demographics are outlined in table 1. During follow-up, until the retention test, none of them had additional surgical, laparoscopic, or simulator experience. Table 2 presents the median scores of the seven novices on the baseline test, the final test and the retention test. It has to be emphasized that a better performance is represented by a lower score. Table 3 compares the novices' median retention test scores with the median experts' scores (performance standard, set as the training goal).

Durability of acquired skills as primary outcome measure

The retention test score did not worsen significantly compared with the final test score for four out of five tasks, pipe cleaner, placing beads, cutting a circle and knot tying (Wilcoxon's signed-rank test). However, deterioration was observed in the score for stretching a rubber band as well as the composed sum score (Table 2).

Table 1. Novices' demographics at retention test.

	Novices (n=7)
Mean age (range) in years	22.7 (21 – 24)
Male n (%)	2 (29)
Median year of study	5 (3 – 5)
Laparoscopic experience	None (except from box training in preceding study)

Table 2. Novices' median scores.

	Baseline test		Final test (I)		Retention test (II)		Difference
Task	Median	(range)	Median	(range)	(I) and (II)	(range)	P-value(*)
Pipe cleaner	333	(126-900)	47	(33-105)	42	(28-63)	0.46
Rubber band	155	(89-484)	49	(22-72)	60	(53-99)	0.03
Beads	831	(474-1558)	235	(168-420)	283	(159-417)	0.24
Cutting Circle	427	(343-520)	134	(89-244)	123	(87-525)	0.31
Knot tying	586	(383-930)	168	(105-223)	182	(60-343)	0.13
Sum score	2631	(2174–2931)	688	(497-971)	800	(515-1219)	0.04

6

Novices' (n=7) performance on three testing moments. Score = time + penalty points. * = Wilcoxon's signed-rank test. A lower score represents a better performance.

Table 3. Novices compared to performance standard.

	Retention test (novice n=7)		Performance standard		Difference
Task	Median	(range)	(expert n=5)	(range)	P-value(*)
Pipe cleaner	42	(28-63)	62	(49-100)	0.06
Rubber band	60	(53-99)	62	(35-195)	0.94
Beads	283	(159-417)	271	(111-318)	0.29
Cutting Circle	123	(87-525)	189	(76-240)	1.00
Knot tying	182	(60-343)	118	(50-177)	0.18
Sum score	800	(515-1219)	705	(351-878)	0.34

Score = time + penalty points. * = Mann-Whitney test. A lower score represents a better performance.

Retentions test results compared with baseline testing and with the performance standard

Comparison of the retention test with the baseline test reveals a significant improvement on all five tasks as well as the sum score. (Wilcoxon's signed-rank test: pipe cleaner, rubber band, placing beads, cutting a circle, knot tying and sum score all $p<.05$) These differences are not displayed in a tabular form. This means that novices perform all tasks faster and more accurately at retention than at baseline testing. No statistical differences were found in any task or the sum score between the novices' retention test and performance standard (Table 3).

DISCUSSION

The previously quantified improvement in laparoscopic skills remained at the same level for four out of five laparoscopic box trainer tasks one year after a basic laparoscopic skills training program. This long-lasting retention of skills is encouraging and supports the implementation of laparoscopic simulator training program at the beginning of residency.

However, deterioration in performance was observed in the task to stretch a rubber band around 16 nails. This finding is remarkable because in that particular task haptic feedback and

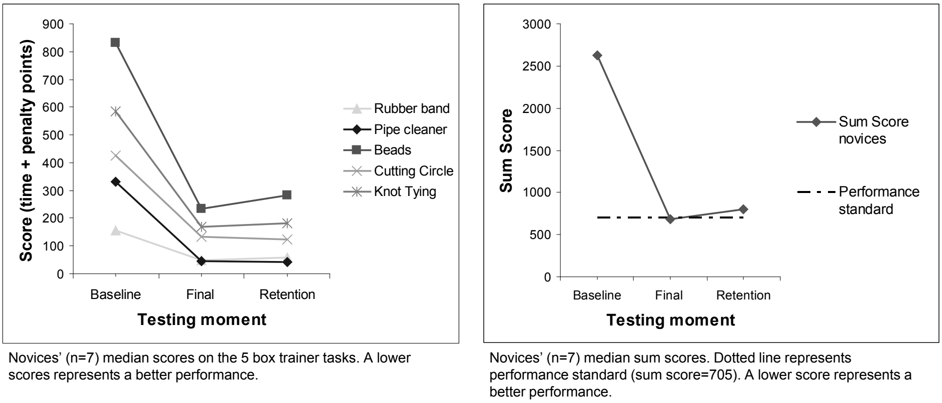


Figure 3. Performance during acquisition and retention.

force transmission play an important role. The ability to adapt to diminished haptic feedback, considered as a substitute for tissue handling, is one of the difficulties of laparoscopic surgery. The current finding may be interpreted as an argument that tissue handling skills are the first to deteriorate in the absence of any practice.

Additionally, the sum score at the retention measurement was worse than the score immediately after the training. It is well possible that the performance in all five tasks slightly decays with time, but that significance was not revealed due to the small number of participants who completed the study. Though, this small decay trend is significant for the composite sum score.

Factors considered to be accountable for the longevity of the acquired skills are the retention interval, the quality of the original training, and trainees' individual differences[Arthur W Jr, 1998; Stefanidis et al., 2006b]. Our 11 months' retention interval is the longest studied so far, as others studied intervals varying between three weeks and seven months. The quality of training is influenced by the type of trainer used for skills acquisition, and the duration, intensity, and goals of the training course. In some studies box trainers were used, in others virtual reality trainers. Stefanidis and colleagues compared both devices and found better skill retention for box trainers[Stefanidis et al., 2005]. In that study, the bean drop task used as in the box trainer mainly requires eye hand-coordination and can be compared to our beads placing task. In general, practice interspersed with periods of rest (*distributed practice*) leads to better acquisition and retention of endoscopic skills than continuous practice (*massed practice*) with little or no rest in between[Moulton et al., 2006; Verdaasdonk et al., 2007b]. Additionally, goal-oriented training leads to consistency of the final results, since all residents are expected to reach the performance standard[Kolkman et al., 2008]. In summary, these data are supportive for the quality of our basic laparoscopic skills course, since we used box trainers, training was held in distributed sessions (one hour weekly for seven weeks), and an expert's performance standard was set as training goal.

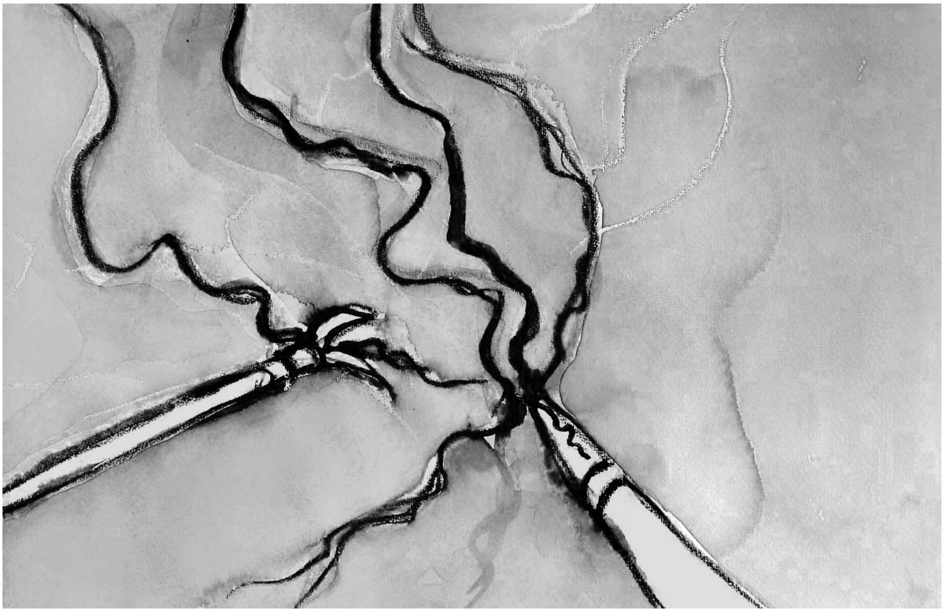
Two striking differences were revealed between our findings and other study results. First, a recent study with a virtual reality trainer revealed that the skills required to perform more difficult tasks deteriorated more than skills needed for the easier tasks 6 months after training[Sinha et al., 2008]. That finding contradicts our finding that the more complex knot tying and cutting skills did not decay significantly, while the placing rubber band task did. Maybe the (relative) resistance to decay of our complex tasks can be explained by a better durability of skills acquired due to the quality of our training, or the small number (seven) of participants failed to show significant deterioration. Second, the study of Vossen and colleagues revealed that most training effect was achieved after 20-30 square knots in a box trainer[Vossen et al., 1997]. This finding contrasts with the small number (seven) of trials needed in our study to achieve the performance standard. However, it has to be noted that the experts in our study only performed one trial of each task for the establishment of this standard. They would probably have shown better performances after they have familiarized with the box trainer and the tasks. The resulting "lower expert level" might be marked as a shortcoming of the study and the value of the result that novices still met the performance standard one year after training may be doubted. On the other hand, this standard revealed to lead to achievable learning goals for skills improvement that sustain over time.

It is of significance to gain insight into the retention of skills in order to realize optimal frequency and efficiency of laparoscopic simulator training. Especially, this is important since students and residents may have a long interval in rotations or residency training before returning to a department in which they can train their laparoscopic skills. Individual differences in retention - and in innate dexterity - among trainees stress the importance of reassessment. Not all subjects may be able to maintain the acquired skill and some require extra training in addition to the training program. To identify these subjects objective skills assessment on the simulator should be performed regularly.

Previously, we have shown that a voluntary training program on a box trainer during residency has a substantial risk to fail[Kolkman et al., 2007b]. Therefore, a goal orientated structured training needs to be implemented into practice in a mandatory fashion, preferably early in residency[Kolkman et al., 2007b]. Specifically, one hour of practice on a box trainer week fits more easily into an already busy residency training schedules than a less efficient training course compressed into two or three days. In order to maintain the acquired skills optimally, it is our opinion that simulator assessment (eventually followed by training) should be repeated at least annually.

In conclusion, our short training program on the box trainer is shown to result in measurable skills improvement that is and merely durable over time. In order to maintain tissue handling skills and to reassure that the skills level for each individual maintains at the performance standard, continuous hands-on practice has to be facilitated and promoted. In order to reach optimal benefit, we recommend the implementation of laparoscopic simulator training program at the beginning of residency training and biannual or annual simulator training and reassessment.

CHAPTER 7
GRADING SURGICAL SKILLS CURRICULA AND TRAINING FACILITIES FOR MINIMALLY INVASIVE SURGERY



Ellen Hiemstra
Henk W.R. Schreuder
Anne M. Stiggelbout
Frank Willem Jansen

Submitted

7

INTRODUCTION

In teaching hospitals all over the world, skills laboratories have been set up in order to train and assess minimally invasive (e.g. laparoscopic) surgical skills outside the operating room in a safe, reproducible environment.[MacRae et al., 2008] This development is driven by quality and patient safety concerns, a restriction in resident working hours and increasing costs of operating room time.[Gould, 2006] Simulator acquired skills are proven to be transferable to the actual operations on patients, leading to a faster operating time and, more important, to fewer errors.[Larsen et al., 2009; Stefanidis et al., 2010]

However, no guideline exists on how to design and use a MIS skills laboratory, nor has a well-recognized standard been defined. The lack of consensus on the appropriate equipment is one of the most common impediments.[Korndorffer, Jr. et al., 2005a] Furthermore, a well-equipped skills laboratory does not automatically generate skilled surgeons. Simulation centres are underutilized, with minimal voluntary use of the models outside the realm of research studies or a structured mandatory training curriculum.[Chang et al., 2007; van Dongen et al., 2008] Nevertheless, there is agreement on at least the need for properly implemented, monitored, and evaluated training curricula for MIS skills.[Korndorffer, Jr. et al., 2006; Park et al., 2002b; Schijven et al., 2008]

This study is an attempt to develop an international and consensus based set of quality criteria for a skills laboratory for training MIS. These criteria include aspects of the design of the skills laboratory and the training curriculum. Quality criteria may help current and future designers and clinicians to implement skills laboratories in their hospitals.

MATERIALS AND METHODS

In order to develop a criteria framework for rating skills laboratories for laparoscopic surgery, the recognised consensus based Delphi approach was used.[Elwyn et al., 2006] This approach enables integrating empirical evidence where it exists with the views of experts.

First, three quality domains were defined: Personnel and resources, Trainee motivation and Curriculum. These domains were inspired on the study of Stefanidis et al. who explored the evidence in the surgical literature regarding laparoscopic curriculum development, and who tried to identify the factors that influence the successful incorporation of simulator training into resident's curriculum.[Stefanidis et al., 2009a] Regarding trainee motivation, *external* motivation of the trainee is addressed, which refers to interventions aimed at modifying behaviour, because the individual *internal* motivation seems difficult to influence.[Stefanidis & Heniford, 2009a]

Additionally, three authors (EH, HS and FWJ) independently searched the current literature for criteria that a skills laboratory should meet and categorized these per domain. For this search, the electronic databases MEDLINE, EMBASE, Current Contents, Science Citation Index and the Cochrane database, were used. In a consensus meeting between the three authors, the lists of criteria were discussed, and an integrated consensus list was formed.

Next, the consensus list was sent electronically to known worldwide experts in training of MIS skills, or a paper version of the consensus list was given to them if he/she visited a congress meeting in 2009 and 2010. An expert in MIS was defined as a gynaecologist, who is well recognized as an expert in advanced laparoscopic surgery, who has three or more publications on MIS related topics, and is actively involved in the organisation of MIS training program in his/her teaching hospital. They were asked to rate each criterion on a 0 to 3 scale in level of importance for a skills laboratory. The definitions of this scale are displayed in table 1. The experts were also instructed to add missing criteria to the list if considered necessary.

Table 1. Definitions rating scale quality criteria

0	not important for rating a skills laboratory
1	optional criterion for a skills laboratory
2	criterion that expresses good quality of a skills laboratory
3	indispensible for a good laboratory

RESULTS

The consensus list contained 9 criteria per domain (Table 2). In total, 23 experts were selected from 14 countries in Europe, North - and South America and Australia. They were either electronically, or in person, asked to fill out the consensus list. All 23 agreed to participate and have rated the nine criteria per domain (Personnel and resources, Trainee motivation and Curriculum). None of the respondents added a new criterion to the list. The results per criterion are displayed as bar charts (Figures 1-3).

In the domain Personnel and resources the presence of a lab technician was considered the least essential for a skills laboratory since it was rated with a median score of 1. The three criteria considered most important were the presence of a curriculum director (laparoscopic expert), the presence of a box trainer and the availability of financial resources. All these criteria received a median score of 3: indispensable for a good laboratory (Figure 1).

In the domain Trainee motivation, the fact that the training should be mandatory is considered the most important. Thereafter, supervision of training by a laparoscopic expert and residents not allowing to perform surgery if the predefined skills level is not reached was considered of importance (Figure 2).

In the domain Curriculum, the presence of over-training facilities (i.e. training after the initially required level of proficiency is achieved) was considered least important (median score: 1). Four criteria were rated with a median score of 3 by the responding experts: the presence of a structured skills curriculum, time dedicated for skills training, maintenance of skills, and a yearly evaluation of the progress in laparoscopic skills of the resident (Figure 3).

As a result, a ranked list of quality criteria is presented, with the ranking based on the median scores of the 23 experts (Table 3).

Table 2. Consensus list

Criterion	Rating			
Personnel & resources	0	1	2	3
1. availability 24 hours a day				
2. space for at least 4 trainees to train simultaneously				
3. presence of a lab technician				
4. presence of a curriculum director (a laparoscopic expert)				
5. presence of a box (/video) trainer				
6. presence of a virtual reality trainer				
7. effective instruction material for the use of the trainer(s) (e.g.video CDrom)				
8. presence of an animal lab				
9. availability financial resources for the skills lab				
Trainee motivation				
1. training sessions are supervised by a laparoscopic expert				
2. training sessions are supervised by a lab technician				
3. a proficiency (i.e. expert) based training goal has been set				
4. the training goal is based on time and precision				
5. training is mandatory				
6. residents are not allowed to perform surgery if predefined skills level is not reached				
7. awards are given for good attendance				
8. presence of tasks of increasing level of difficulty				
9. variability is present in the laparoscopic tasks				
Curriculum				
1. presence of a structured skills curriculum				
2. time is dedicated for skills training in the residency curriculum				
3. monthly training sessions are organized				
4. presence of "over training" (i.e. better than training goal) facilities				
5. Repetitive training over various training sessions				
6. Maintenance of training				
7. Retention of skills is established every 12 months				
8. training goal increases with progression in residency				
9. progress in laparoscopic skills is incorporated in yearly evaluation of resident				

DISCUSSION

For the setting of a laparoscopic skills laboratory in a (teaching) hospital the bottom line is that a box trainer model and financial resources are required. The training has to be mandatory, to be supervised by a laparoscopic expert and residents should not perform (supervised) in vivo laparoscopic surgery if the predefined skills level is not reached. Skills training should be imbedded in a structured curriculum with time scheduled for training. Finally, maintenance of skills, and a yearly evaluation of the skills level are recommended. Our detailed consensus list can be used when a MIS skills laboratory has to be set. Furthermore, it gives cues for verifying

Table 3. Ranked list of quality criteria (median scores of 23 experts).

Criterion	Median score (range)
Personnel & resources	
1. availability 24 hours a day	2 (0-3)
2. space for at least 4 trainees to train simultaneously	2 (0-3)
3. presence of a lab technician	1 (0-3)
4. presence of a curriculum director (a laparoscopic expert)	3 (1-3)
5. presence of a box (/video) trainer	3 (2-3)
6. presence of a virtual reality trainer	2 (0-3)
7. effective instruction material for the use of the trainer(s) (e.g.video)	2 (0-3)
8. presence of an animal lab	2 (0-3)
9. availability financial resources for the skills lab	3 (2-3)
Trainee motivation	
1. training sessions are supervised by a laparoscopic expert	3 (2-3)
2. training sessions are supervised by a lab technician	2 (0-3)
3. a proficiency (i.e. expert) based training goal has been set	2 (0-3)
4. the training goal is based on time and precision	2 (0-3)
5. training is mandatory	3 (1-3)
6. residents are not allowed to perform surgery if predefined skills level is not reached	3 (1-3)
7. awards are given for good attendance	2 (0-3)
8. presence of tasks of increasing level of difficulty	2 (0-3)
9. variability is present in the laparoscopic tasks	2 (1-3)
Curriculum	
1. presence of a structured skills curriculum	3 (2-3)
2. time is dedicated for skills training in the residency curriculum	3 (1-3)
3. monthly training sessions are organized	2 (0-3)
4. presence of "over training" (i.e. better than training goal) facilities	1 (0-3)
5. Repetitive training over various training sessions	2 (1-3)
6. Maintenance of training	3 (1-3)
7. Retention of skills is established every 12 months	2 (0-3)
8. training goal increases with progression in residency	2 (1-3)
9. progress in laparoscopic skills is incorporated in yearly evaluation of resident	3 (1-3)

the quality of an already existing laboratory, just by using the list of ranked quality criteria as a checklist. From there, the focus for improvement or new developments can be chosen.

In the domain Personnel and resources, the presence of a box trainer is considered relatively more important than the presence of a virtual reality (VR) trainer. This finding is consistent with recent results of Palter et al, who found in their inventory that residents prefer box trainers above VR simulators for training the more advanced laparoscopic skills.[Palter et al., 2010] On the contrary, both trainer models have a good correlation for the assessment of laparoscopic skills.[Newmark et al., 2007] VR trainers have the advantage they allow solitary

7

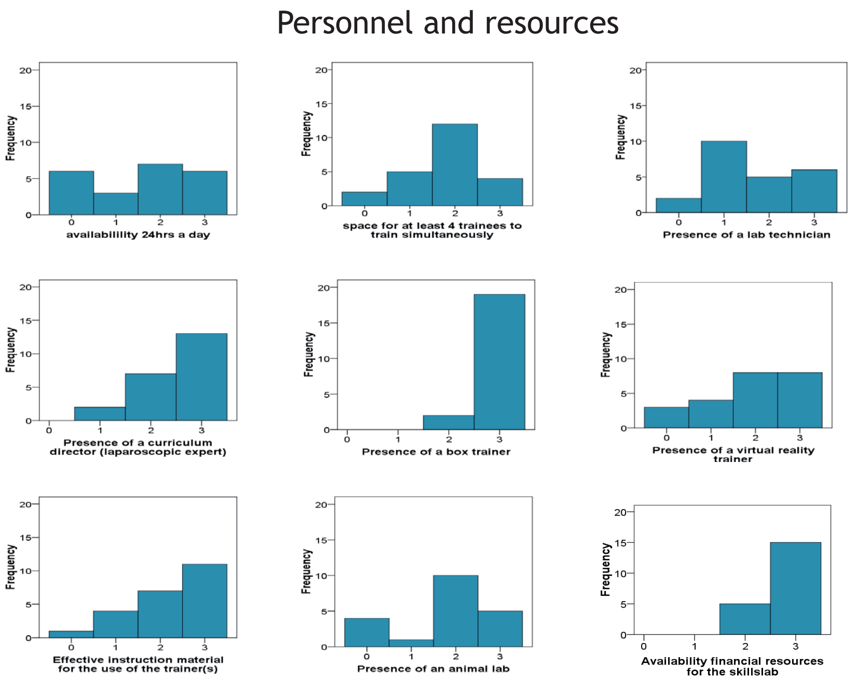


Figure 1. Expert opinion domain 'Personnel and resources'.

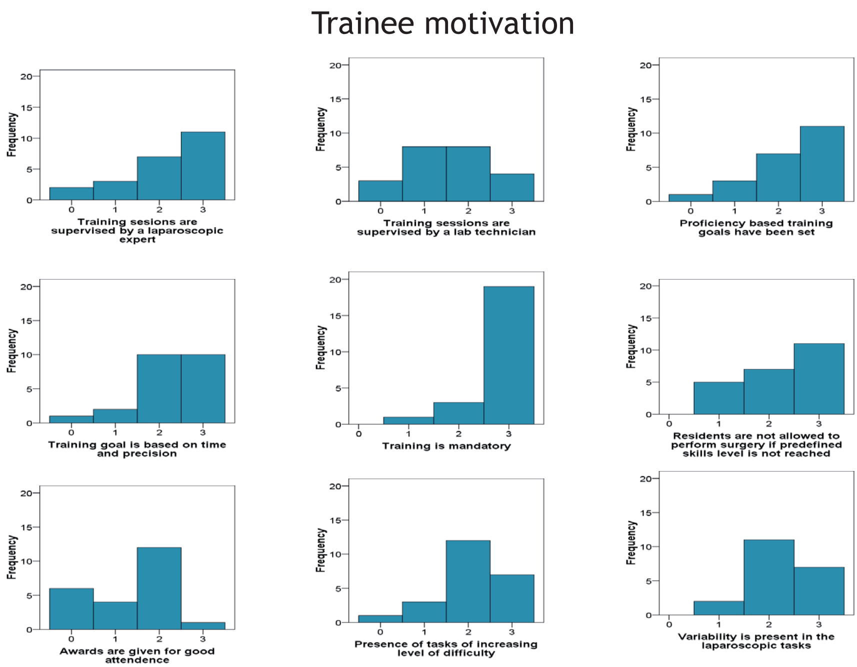


Figure 2. Expert opinion domain 'Trainee motivation'.

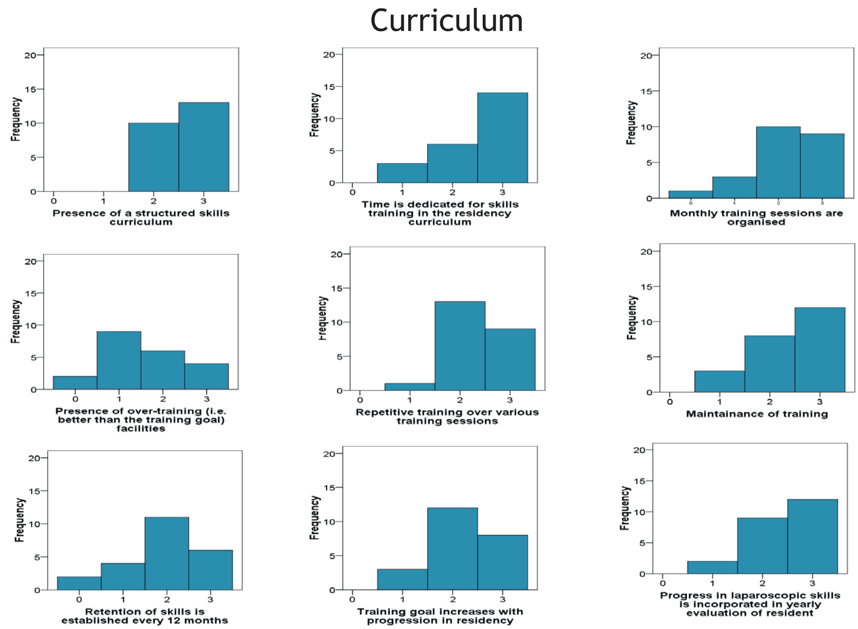


Figure 3. Expert opinion domain 'Curriculum'.

training while the supervisor can monitor the resident's skills level electronically. On the other hand, the presence of a supervisor required during box training has the advantage that surgical knowledge can be transmitted. Furthermore, the presence of a laboratory technician is rated low. This could be explained by the fact that an enthusiastic laparoscopic expert can fulfil this role. However, in our opinion the presence of a permanent availability of a technician gives a professionalizing of skills laboratory, with all its advantages.

In parallel with the importance of setting the training mandatory, it was found that most residents do not reach the performance standards of basic laparoscopic skills if the skills training is voluntary.[Kolkman et al., 2005] Furthermore, training up till a predefined level of skills is superior over training based on the time spent. In fact, the time required varies and training till a certain level induces an external motivation. Ideally the training should be proficiency based[Korndorffer, Jr. et al., 2005c] and supervised by a laparoscopic expert. Training exercises should not be based on time only, and a score for precision should be added.[Smith et al., 2002] It can be argued whether the exercises should have an increasing level of difficulty. On the one hand this may keep the trainees motivated throughout their entire specialty training, on the other hand, basic laparoscopic skills should be acquired as early as possible in residency after which residents can expand their proficiency in the operating room in learning anatomy, pathology, and operating techniques, while maintenance of the basic skills is all there is left to do.[Korndorffer, Jr. et al., 2005c]

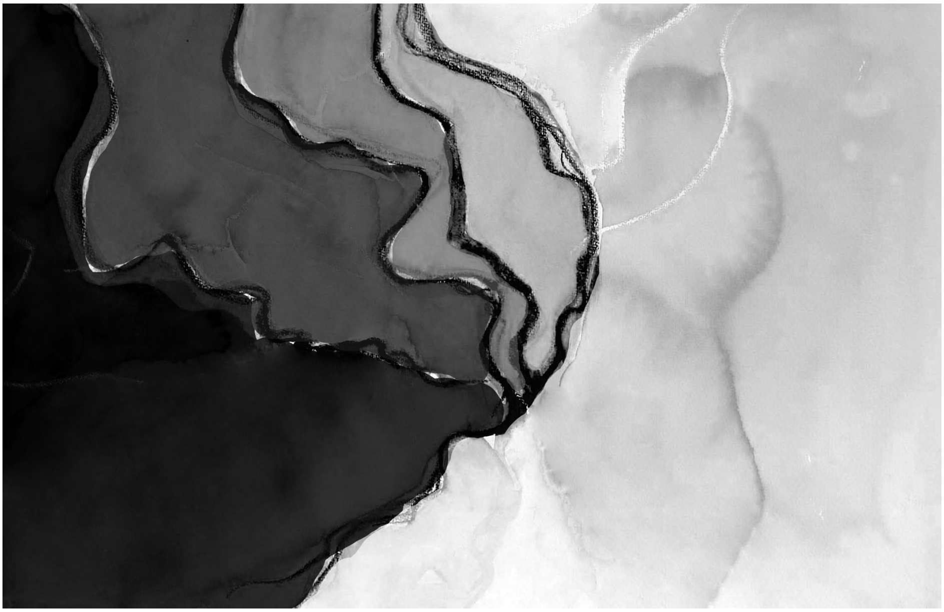
In the third domain Curriculum there is a clear consensus about incorporating the skills training for MIS in a proficiency based training curriculum. It is important to dedicate time for skills training during working hours and organize repetitive training sessions. Overall, the presence of a mandatory, structured and competency based skills training curriculum is the key to success.[Gallagher et al., 2005; McClusky, III et al., 2008; Stefanidis & Heniford, 2009a]

With the increasing pressure on guaranteed skilfulness of surgeons, many MIS specialty teaching hospitals feel the need to implement training facilities outside the OR. Although it is essential to define the purpose and to identify resources early in the development of a skills laboratory, the reality is often the other way around.[MacRae et al., 2008] As a result, many hospitals have designed laboratories based on an individual trainer's ideas and preferences. Besides, curriculum development is lagging.[Stefanidis et al., 2010] The strength of this study is that a consensus based rating system has been developed with agreement of laparoscopic experts all over the world. However, the selection of the 23 experts might be a limiting factor, because it depended on the definition we chose which was in part based on their reputation in their peers field.

A generally accepted set of criteria potentiates a system of accreditation for laparoscopic skills laboratories. Similarly, the American College of Surgeons has developed a system for accreditation of laboratories regarding general surgical skills in institutes.[2009b] These criteria are used to determine whether an institute meets the minimum requirements for accreditation as a Level II (Basic Education) or a Level I (Comprehensive Education) institute. In parallel, our set of criteria can be used as a framework useful in daily practice and possibly for accreditation purposes in the future. More in detail, a skills laboratory can be assessed rating the presence of a criterion with the corresponding median score of our ranked quality criteria list. That way, criteria that are considered more relevant according to our expert panel receive higher ratings. As a result, a MIS skills laboratory with a MIS skills curriculum can obtain at maximum 62 points (20 points for Personnel and resources, 21 points for Trainee motivation and 21 points for Curriculum). This total score can be used to choose the focus for future developments. Additionally, a practical application might be that a basic MIS laboratory should have at least the criteria with a median score 3, while a comprehensive MIS laboratory should also have all criteria with a median score of 2 for the certification.

In conclusion, this rating list can be used to set up and maintain a minimally invasive skills laboratory. In a skills laboratory, at least a box trainer has to be present with a proficiency based training program. The training should be incorporated in a formal curriculum which is obliged prior to attendance of real in vivo surgery in order to enhance patien safety.

CHAPTER 8
THE VALUE OF AN OBJECTIVE ASSESSMENT TOOL IN THE OPERATING ROOM



Ellen Hiemstra
Wendela Kolkman
Ron Wolterbeek
J. Baptist M.Z. Trimbos
Frank Willem Jansen

Adapted from Can J Surg, 2011; 54: 116-22

8

INTRODUCTION

Nowadays, it is becoming more and more difficult to achieve surgical proficiency. Residents experience less training due to reduced working hours and a decreased surgical caseload.[Hammond et al., 2006] Additionally, with the development of new surgical techniques, skills acquisition is more challenging.[Haluck & Krummel, 2000] Currently, basic surgical procedures are sufficiently mastered after finishing residency training, but advanced procedures are not.[Kolkman et al., 2006] Ultimately, skills deficiencies will impede post-residency performance.[Shay et al., 2002] Moreover, residency programs still rely heavily on informal and subjective evaluations based on recollections of supervisors.[Kolkman et al., 2005; Mandel et al., 2000] Therefore, on one hand, surgical skills training need to become more efficient, and on the other hand, appropriate assessment is required in order to optimally benefit from the spare learning moments in the operating room (OR).

An objective assessment tool can fulfil an important role during operative training.[Aggarwal et al., 2008; Beard, 2007] Such a tool can be an aid to the learning process through constructive feedback on performance. Secondly, an assessment tool can be applied to establish competency levels and to mark progression in time. Finally, it can provide a benchmark criteria to be used as a training goal or for credentialing purposes.[Cuschieri et al., 2001; Darzi et al., 1999]

To fulfil this need for an objective assessment tool, the OSATS (Objective Structured Assessment of Technical Skills) was developed by Martin et al. in Toronto in 1997.[Martin et al., 1997] An OSATS consists of a procedure-specific checklist, a pass/fail judgment and a global rating scale. The latter turned out to be superior in terms of reliability and validity.[Goff et al., 2000; Martin et al., 1997; Swift et al., 2006] On this global rating scale, domains are scored on a 1 to 5 Likert-scale, with an explicit description at point 1,3 and 5.

So far, studies about the quality of OSATS have mainly been conducted in simulators or live animal models.[Reznick et al., 2006] Although applying OSATS in simulator settings has the benefit that repeated practice is enabled without the risk to harm patients, simulators will never perfectly mimic operative conditions. Therefore, OSATS have been implemented for the assessment of real surgical procedures on a large scale in the Netherlands in residency programs. Moreover, plans are being developed to use this form of assessment tool for certification purposes after residency training. However, only a few studies have investigated the value of intraoperative use of OSATS.[Aggarwal et al., 2008; Bodle et al., 2008] Aggarwal et al. found that the OSATS score discriminates between a novice and an expert surgeon performing a laparoscopic cholecystectomy demonstrated by video-based assessment.[Aggarwal et al., 2008] Bodle et al. concluded from feedback questionnaires that trainers and trainees in the United Kingdom perceived the OSATS to be valid and valuable.[Bodle et al., 2008] In the absence of data on the implementation of OSATS in daily practice, the current study was conducted in order to assess its value in clinical practice by analysing residents' learning curves for a variety of surgical procedures in gynaecology.

MATERIALS AND METHODS

In the Netherlands, the Obstetrics and Gynaecology (Ob/Gyn) residency program lasts six years. On average, three of these six years are spent in a university teaching hospital, and the complementary period is spent in a non-university teaching hospital. The university hospitals provide a curriculum to train residents in a variety of subspecialties, like reproductive health care, perinatology and oncology. Specifically, a three-month clinical rotation is spent on gynaecological surgery. During this rotation, which is generally attended during the fourth postgraduate year (PGY), residents are scheduled to perform surgery in the OR for four days a week. Gradually, a resident is given more responsibility as experience accrues, depending on the resident's technical skills, the type of procedure and patient characteristics. Finally, a resident performs a procedure as the primary surgeon, in the presence of a supervising consultant.

Study Design

In 2005, the global rating scale of the OSATS (referred to as "OSATS" in this thesis) was introduced at the department of Ob/Gyn of the Leiden University Medical Center in an observational study of its implementation in clinical practice (Figure 1). The assessment tool had been adapted from Martin et al.[Martin et al., 1997] The six domains of an OSATS represent aspects of technical competence in surgery. The only modification to the original form is that we merged the domains 'knowledge of instruments' and 'instrument handling'. This is in accordance with the version of the OSATS form used by the Royal College of Obstetrics and Gynaecology.[RCOG

OSATS - global rating scale of operative performance

Please circle the number corresponding to the candidate's performance in each category, irrespective of training level.

	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 and Motion:	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economy of movement and maximum efficiency
Knowledge and handling of instrument:	Lack of Knowledge of Instruments		Competent use of instruments but occasionally appeared stiff or awkward		Obvious familiarity with instruments
Flow of operation:	Frequently stopped procedure and seemed unsure of next move		Demonstrated some forward planning with reasonable progression of procedure		Obviously planned course of procedure with effortless flow from one movement to the next
Use of assistants:	Consistently placed assistants poorly or failed to use assistants		Appropriate use of assistants most of the time		Strategically used assistants to the best advantage at all times
Knowledge of specific procedure:	Deficient knowledge. Needed specific instructions at most steps		Knew all important steps of procedure		Demonstrated familiarity with all aspect of operation

Figure 1. OSATS form.

2009] During this implementation study, residents were instructed to register an OSATS assessment of every procedure that they performed as a primary surgeon during their three-month rotation in gynaecological surgery. Procedures during which a resident independently performed some important steps were included as well. After the supervising consultant had filled out the OSATS form, the results were discussed with the resident in order to provide him/her with constructive feedback per domain.

While the assessed trainees were PGY 4 Ob/Gyn residents, the supervisors could be any gynaecologist working as a consultant at the department who was present supervising the surgical procedure. They were instructed how to complete the OSATS form. In essence, the instruction was to mark the number on the Likert-scale corresponding to the resident's performance on each domain, irrespective of the training level.

Individual learning curves

All OSATS were collected, and data were analysed using an SPSS-program for Windows (SPSS version 16.0 SPSS Inc., Chicago, IL). The total score of each OSATS was calculated by adding up the score of the six domains (at minimum 6 and maximally 30 points). An OSATS score of 24 points equals the score in which each domain at average is rated with 4 points (75% of the maximally score that ranges from 1 to 5). This score was chosen as a threshold for good surgical performance, in the absence of benchmark criteria in other studies. Learning curves for each individual resident were drawn by plotting his/her OSATS scores against the total caseload during a clinical rotation, regardless of which procedures were performed. To establish the caseload, all consecutively performed procedures that were assessed with an OSATS were numbered. For each resident, the mean OSATS score during the rotation was calculated, and progression in time was illustrated by mapping a regression line.

Construct validity

No 'gold standard' is available to measure surgical performance. Therefore, the construct validity (i.e. the extent to which a test measures the trait that it purports to measure) should be used to verify the quality of an assessment tool for surgical skills.[Feldman et al., 2004a; Moorthy et al., 2003] In this study, the construct validity of OSATS was established by testing the hypothesis that surgical performance improves as the procedure-specific experience accrues. For that purpose, the average learning curve for the 'average' procedure was mapped by plotting the OSATS score against the procedure-specific caseload. The procedure-specific caseload was also based on the number of assessed procedures.

To test this hypothesis, a linear relation between OSATS score and experience was assumed. The advantage of simplifying the average procedure-specific learning curve to a straight line is that the performance level at the start can be determined, as well as the amount of progression in technical surgical skills, taking individual performance levels and learning potential into account. Therefore, a linear mixed model was fitted as random coefficients model with a random slope and a random intercept per resident. P-values <.05 were considered statistically significant, and ninety-five per cent confidence intervals (95% CI) were calculated.

Objectivity of assessment with OSATS

After this implementation study, the opinion of assessed trainees and supervisors was questioned regarding the objectivity of an assessment with an OSATS. They were asked to rate the OSATS on a Likert-scale ranging from 1 "subjective" to 5 "objective". The assessed trainees were residents who were recruited during an education afternoon in the LUMC of which the attendance was obligatory during Ob/Gyn residency training. The supervisors were the same consultants who had participated in the implementation study.

RESULTS

Nine residents attended a three-month clinical rotation in gynaecological surgery, and agreed to participate in the study. Three were male and six were female. Nineteen different types of procedures were assessed with an OSATS, and the total number of procedures was 319. Among these procedures, 39% were abdominal, 31% were laparoscopic, and 20% were procedure with a vaginal approach, and the remaining 10% were hysteroscopies (Table 1). On an individual basis, the median number of procedures assessed was 40 (range 12-60).

Individual learning curves

The nine individual learning curves were drawn by plotting OSATS scores against the total caseload (regardless of which specific procedure had been performed) during the clinical rotation (Figure 2). The regression lines of these curves are displayed too, together with the threshold of 24 (out of 30) OSATS points. Regression analysis revealed that the two residents with the lowest average scores (resident A and B) did not reach the threshold of 24 points within their clinical rotation. Resident C and D reached the threshold while nearing the end of their rotation. Only resident H and I achieved relatively high scores at the start of the three-month period and continued to show improvement.

Average procedure-specific learning curve

Additionally, the average OSATS scores were plotted against the experience, i.e. the procedure-specific caseload, for the first ten procedures (Figure 3). The resulting average learning curve within procedure passed the threshold of 24 points at a caseload of five procedures. Additionally, a plateau in performance was reached after a caseload of eight procedures. To establish the construct validity of OSATS it was tested whether the OSATS score increased significantly with an increasing caseload using a linear mixed model. The slope of the general learning curve was 1.10 OSATS points per assessed procedure ($p<.01$, 95% CI: 0.44 – 1.77). In other words, the average performance based on total OSATS score improved by 1.10 points for every consecutively performed procedure.

An OSATS score of 24 was set as the performance standard. The dotted line is based on linear mixed model analysis.

Objectivity of the assessment

The supervisors were 21 gynaecologists, all working as consultants at the Department of Gynaecology at the LUMC. The median OSATS score given to residents by each supervisor

Table 1.

Procedures	Number assessed with an OSATS
Laparoscopic procedures	98
Diagnostic laparoscopy or sterilization	23
(Bilateral) Salpingo-oophorectomy	41
Cystectomy	17
Ectopic Pregnancy (tobotomy or tubectomy)	4
Total laparoscopic hysterectomy	13
Hysteroscopic procedures	31
Diagnostic hysteroscopy	12
Therapeutic hysteroscopy	19
Abdominal procedures	125
Abdominal hysterectomy (with (B)SO)	42
Resection myoma, endometrioma or adnexectomy	6
Caesarean section	64
(Interval) debulking	7
Sacrocolpopexy	5
Procedure with a vaginal approach	65
Vaginal hysterectomy	43
Anterior and/or posterior colporrhaphy	6
(Partial) vulvectomy	2
Operation of cervix (cerclage or conization)	8
Anal sphincter repair	4
Laser treatment vulva	2
Labioplasty	1
Total	319

ranged from 18 to 30, and the number of assessed procedures ranged from 1 to 114. Moreover, some gynaecologists assessed only one specific procedure (e.g., a caesarean section), while others assessed the entire surgical spectrum.

All 24 residents who were present at the obligatory education afternoon answered the question about the OSATS. One person was excluded from analysis due to inexperience with this assessment form because of being just at the start of residency training. Residents rated the OSATS with a median score of 2 (range 1-4 on a 5-point Likert scale with 1:subjective to 5:objective). The median score of the supervisors was 3 (range 1-4).

DISCUSSION

Intraoperative OSATS can be used to assess resident's surgical training over time. By plotting the OSATS score against experience it can be determined whether, and how much, progression is

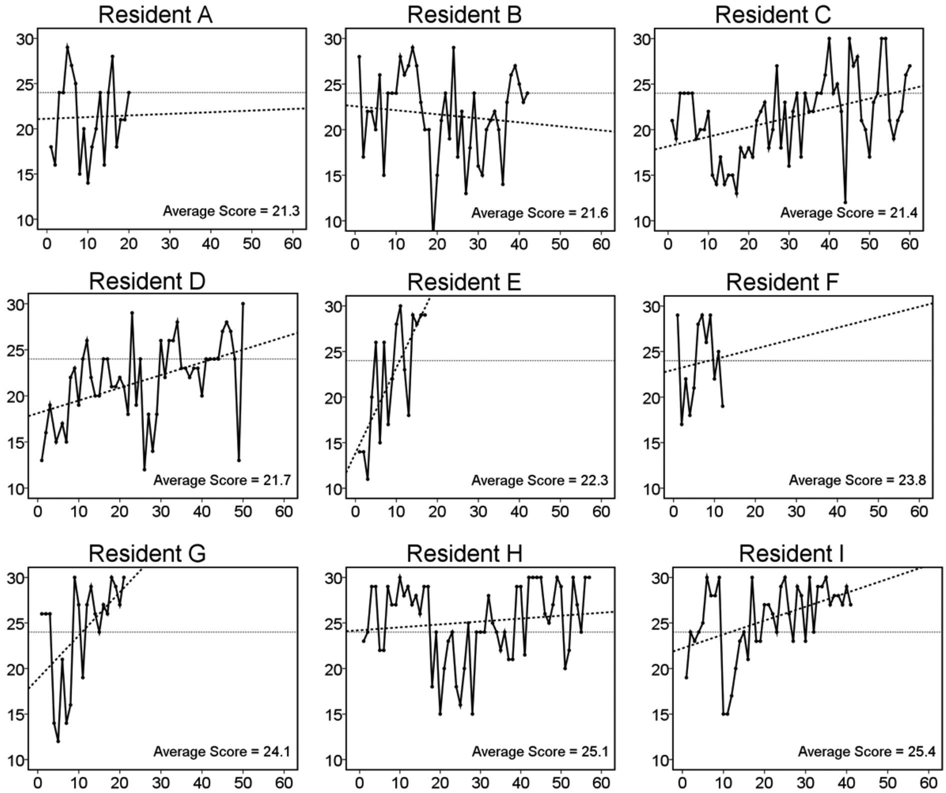


Figure 2. Individual learning curves (regardless the type of procedure performed). x = total caseload expressed in number of assessed procedure (regardless of the type of procedure performed), y = performance expressed in total OSATS score, dotted line = individual regression line.

present. The use of an objective assessment tool is a new way to establish learning curves. Prior parameters are the operation time, the complication rate and the conversion rate in case of laparoscopic procedures.[Altgassen et al., 2004; Kolkman et al., 2007a] However, operation time and complication rate have shown to be crude and indirect as these largely depend on the difficulty of the individual surgical case (e.g. the co-morbidity of a patient), and the supervising surgeon.[Moorthy et al., 2003] The intraoperative use of OSATS may overcome these disadvantages.

Two out of nine residents did not progress beyond the benchmark level of 24/30 OSATS points within the three-month clinical rotation. This failure is likely to be a sign of stagnation of their learning process, and can only partially be explained by the coincidence that they encountered more complex procedures later in their rotation. Additionally, only two residents showed good performance during the entire clinical rotation, taking the average OSATS scores and the progression into account. This small proportion illustrates the concern whether current residency programs with work hour restrictions sufficiently fulfil the need to master surgical proficiency.

Secondly, the construct validity of the OSATS for assessment purposes was revealed by confirming the hypothesis (i.e. the construct) to be true that a resident's OSATS' score

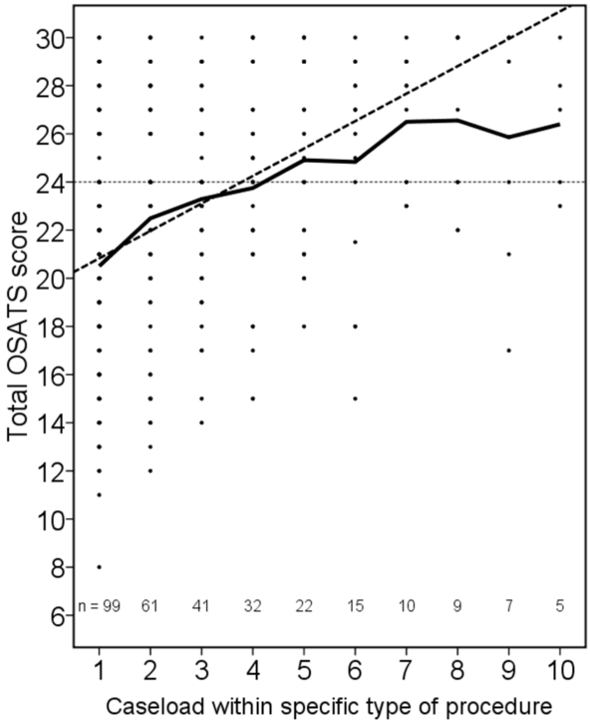


Figure 3. Average Objective Structured Assessment of Technical Skills (OSATS) scores plotted against procedure-specific learning curve for the first 10 procedures.

improves as procedure-specific experience accrues. This is not the conventional way to prove the construct validity. However, it is a more subtle approach than the often used method to confirm the ability of an assessment tool to discriminate between two groups of hugely varying level of experience. That was done by Aggarwal et al. who revealed that experienced surgeons have higher OSATS scores than novice surgeons for one standardized procedure, the laparoscopic cholecystectomy.[Aggarwal et al., 2008] The straight line model we used as an argument for the construct validity has two limitations. Surgical performance cannot infinitely improve (the maximum OSATS score is 30 points), and secondly, the learning curve for surgical skills consists of an initial steep phase, then changes slowly until the curve becomes more flat.[Dagash et al., 2003] However, the advantage of simplifying resident's learning curve to a straight line, and additional analysis with linear mixed model, is that progression in surgical skills can be quantified taking the individual level of performance and learning potential into account. From this data, it was found that a resident's performance improves with 1.14 OSATS points at average every time the same procedure is performed (and assessed). Of course, we may not simply generalize this conclusion, because this increase is the average of 19 very different surgical procedures.

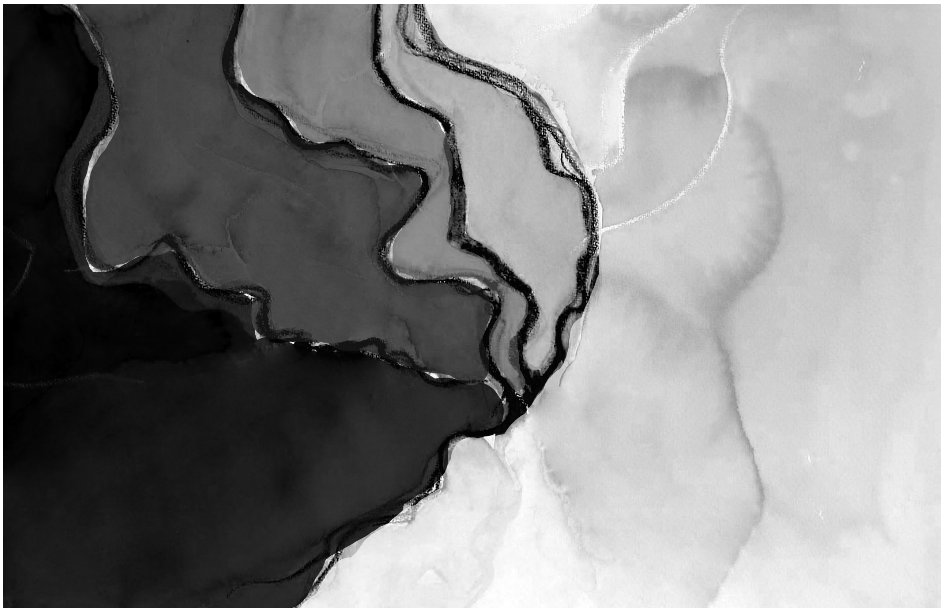
The aforementioned formation of a plateau in the real situation is observed in the average procedure-specific learning curve. This plateau is achieved after a caseload of eight (of the same) procedures. This is in accordance with results of a questionnaire held among residents is which they judged a number of ten of the same procedures necessary to be a safe and confident surgeon.[Rattner et al., 2001] Again, this value of this generalization is limited because of the heterogeneous range of assessed procedures.

This study was conducted under regular clinical conditions. Therefore, even the same procedures widely varied with respect to difficulty and complication risk. Also, variation shall have been present in the extent to which consultants allowed residents to independently perform a surgical procedure. Furthermore, the assessment rate might not be 100 per cent. The resulting selection bias may be in favour of the best performed procedures. However, not all procedures need to be assessed to gain insight in the progression of an individual resident. More importantly, the intended objectivity of assessment with an OSATS seems to be disappointing, taking the finding that none of the residents, nor any staff member, valued the OSATS to be objective into account. Additionally, the number of assessed procedures and the OSATS-score varied enormously among the consultants. This variation occurred despite the uniform instruction that all supervisors had received. An attempt to achieve more uniformity might be realized by organizing additional training for the supervisors in the registration of an OSATS. However, in our opinion, the effect of such training is limited. No information can be added to the original instruction to mark the number on the rating scale corresponding to the resident's performance on each domain, irrespective of the training level. Moreover, an assessment based on the opinion of an individual will never be free from subjectivity. A study in which residents all perform at least ten of the same procedures consecutively would have allowed firmer conclusions about the learning for curve of that specific procedure. However, insight in daily practice is obtained by analysing the heterogeneous data of our study, and illustrates the study's relevance.

In conclusion, assessment with OSATS during residency has many advantages. OSATS-based learning curves have the potential to select residents in need of more guidance during their learning process. Consequently, cues are provided to tailor surgical skills training to individual needs. An OSATS does not need to concern the entire procedure; (small) steps of the procedure can be evaluated as well. Additionally, it provides a framework of structured instantaneous feedback on surgical skills in general (total OSATS score). Theoretically, the specific domains of technical skills (e.g. respect for tissue, knowledge and handling of instruments) also provide cues for identifying individual needs. However, the information that the domain-specific scores add is limited as revealed from the small variety of score within one OSATS. Ideally, the structural feedback on surgical performance using assessment with OSATS will enhance the efficiency of the spare learning moments in the OR. From that point of view, we consider the general global rating scale of OSATS to be suitable for large scale implementation in the OR.

However, the inherent subjectivity of an assessment using an opinion based tool needs to be taken into account. Regarding the results of the questionnaire and the enormous variation in supervisor's scores, an OSATS unfortunately is not as objective as it intended to be. This is an important limitation of the OSATS that, to our knowledge, has not been highlighted in other publications about this assessment tool. Furthermore, there are other ways to evaluate a subject's surgical skills. Therefore, caution needs to be exercised in using OSATS for certification and qualification purposes, or in advising an individual resident to choose for a non-surgical specialization if the OSATS-based performance continues to be disappointing. Though, presently, it seems to be the best tool available.

CHAPTER 9
IMPLEMENTATION OF OSATS IN THE RESIDENCY PROGRAM: A BENCHMARK STUDY



Ellen Hiemstra
Navid Hossein pour Khaledian
J. Baptist M.Z. Trimbos
Frank Willem Jansen

Submitted

INTRODUCTION

The exposure to surgical procedures during residency training has decreased. Consequently, residents have to achieve the same competencies in fewer working hours than their counterparts some decades ago.[Haluck & Krummel, 2000] The decrease in working hours is exaggerated by a trend to non-surgical therapies for certain traits.[Hammond & Karthigasu, 2006] With this development, the apprenticeship model relying on experiential training, large number of procedures performed and subjective, observational assessment of surgical skills no longer suffices. Instead, a different and highly efficient program is required to achieve surgical proficiency.

However, a large survey in 1998 in the United States revealed that evaluation of surgical skills is usually done by subjective faculty assessment at the end of a rotation, and, therefore, is based on recollection of events over time.[Mandel et al., 2000] Obviously, this method lacks reliability and validity. Therefore, there is a pressing need to evaluate surgical skills more objectively in a structured fashion.

In response, assessment tools have been developed to evaluate surgical skills.[Reznick et al., 1997] An example of such a tool is the OSATS (Objective Structured Assessment of Technical Skills). OSATS assesses discrete domains of surgical competence. It has proven to be reliable and construct-valid in bench models and in live animal models.[Martin et al., 1997] Goff et al. showed the reliability and validity of OSATS in animal models for residents in Obstetrics and Gynaecology.[Goff et al., 2000]

Training on bench models in skills laboratories enables repeated skills training without the risk of harming patients. However, practising on a bench model is not equivalent to performing surgery on a living patient in the operating room (OR). Obviously, a surgeon has to become proficient in the latter. As a consequence, scientific evidence of more objective intraoperative assessment tools is needed. In fact, the construct validity of intraoperative use of OSATS has been proven in two studies[Aggarwal et al., 2008; Hiemstra et al., 2011]. Additionally, residents and supervisors perceive intraoperative administration of OSATS to be a valuable and valid tool, as was revealed from a questionnaire in United Kingdom.[Bodle et al., 2008]

However, prior to large-scale implementation of assessment with OSATS, more information is required.[Bodle et al., 2008] It has been stated that an objective assessment tool can be used for authorization,[Darzi et al., 1999] but cut-off values have not been defined.[van Hove et al., 2010] Additionally, it is unclear whether residents' self-assessment is in accordance with their supervisor's rating using the OSATS. This can be interpreted as a form of inter-rater reliability, and is important in the renewal of residency programs with an increasing focus on self-assessment. Finally, we are interested in the opinion of clinicians who will have to work with intraoperative administration of the OSATS after its implementation.

The aim of our prospectively designed study is threefold; first, to establish at which OSATS score a supervisor judges a resident able to perform the procedure autonomously; second, to evaluate the reliability between resident and supervising gynaecologist regarding intraoperative assessment of technical surgical skills, and finally, to question aspects of the satisfaction of residents and supervisors with the intraoperative administration of OSATS.

MATERIALS AND METHODS

To answer the research questions, study-specific OSATS forms were distributed for application in clinical practice. Additionally, a survey was performed among users of the OSATS.

Intraoperative administration of OSATS

All obstetrics and gynaecology (Ob/Gyn) residents who were attending their specialist training at the Leiden University Medical Center (LUMC) or at their affiliated teaching hospitals from July 2007 to June 2009, were asked to participate in the study. They were informed by means of mailing and by an individual briefing. Residents were asked to complete a study-specific double-sided assessment form after each procedure they had performed as the primary surgeon. One side had to be filled out by the resident, and the other side by the supervising consultant. Each side of the form contained a general global rating scale of the OSATS and the question whether the resident would have been able to perform the procedure autonomously, i.e. without supervision (Figure 1). On the global rating scale, which had been adapted from Martin et al.[Martin et al., 1997], six domains of technical surgical skills could be rated on a 1 to 5 scale in which ratings 1, 3 and 5 had an explicit description. The ability to perform a certain procedure autonomously was rated on a 3-points scale (no, maybe or yes).

After the self-assessment on one side of the form, the supervisor had to complete the other side blind for the results of the self-assessment. Next, the results were discussed with

OSATS – global rating scale of operative performance

Please circle the number corresponding to the candidate's performance in each category, irrespective of training level.

	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 and Motion:	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economy of movement and maximum efficiency
Knowledge and handling of instrument:	Lack of Knowledge of Instruments		Competent use of instruments but occasionally appeared stiff or awkward		Obvious familiarity with instruments
Flow of operation:	Frequently stopped procedure and seemed unsure of next move		Demonstrated some forward planning with reasonable progression of procedure		Obviously planned course of procedure with effortless flow from one movement to the next
Use of assistants:	Consistently placed assistants poorly or failed to use assistants		Appropriate use of assistants most of the time		Strategically used assistants to the best advantage at all times
Knowledge of specific procedure:	Deficient knowledge. Needed specific instructions at most steps		Knew all important steps of procedure		Demonstrated familiarity with all aspect of operation

Is the resident able to perform the procedure autonomously?	No Maybe Yes Please circle the most appropriate answer

Figure 1. OSATS form used for the study.

the resident in order to create a learning opportunity. The supervisor could be any consultant gynaecologist working as a staff member at one of the hospitals of the study. They had uniformly been instructed to fill out the OSATS form by rating the resident as objectively as possible, irrespective of his/her level of experience.

Questionnaire

Additionally, a survey was held among the users of the OSATS in July 2009, regarding the user's satisfaction with the intraoperative assessment tool. This questionnaire was sent to residents who were attending their Ob/Gyn residency training at that time and to the consultant gynaecologists who actively participated in teaching surgical skills in the ORs of these hospitals. A Likert scale was used to have the respondents express their agreement or disagreement with five statements on OSATS on a five-point scale. The five statements included that OSATS is a valid instrument, it is subjective (or objective), it should be used for assessment, it helps in acquiring surgical skills, and it leads to irrelevant paperwork. Finally, the participants were asked for their opinion on the ideal frequency of administration of an OSATS.

Statistics

The results were collected in the statistical SPSS program (SPSS, version 16, SPSS inc., Chicago, IL). The corresponding median OSATS scores were calculated for a positive, uncertain and negative response to the question of whether a resident was able to perform the procedure autonomously. The intraclass correlation coefficient (ICC) for single measurements was used to determine the inter-observer reliability. Although arbitrary, a frequently used nomenclature for the ICC is that a score ranging from 0.41 to 0.60 indicates moderate agreement, a score ranging from 0.61 to 0.80 substantial agreement, and a score above 0.81 indicates perfect agreement. [Landis et al., 1977] However, the use of ICC can be deceptive, as the outcome is also dependant of the number of items scored. However, we have chosen this measure to enable a comparison with results of other studies about the reliability of the OSATS.

RESULTS

The participants were 19 residents, equally distributed among all six postgraduate years. Six were male, thirteen were female. In total, 127 study forms were collected and the data analysed. The procedures assessed related to abdominal, vaginal, laparoscopic and hysteroscopic surgery. For an individual resident, the median number of procedures assessed was 14 (range 1-28). A total of 27 gynaecologists assessed these procedures.

The OSATS score had been completed by the gynaecologist in all the forms returned (100%). In 122 cases (96%), the question about the ability to perform the procedure autonomously was also completed. The OSATS score on the resident's site was filled out in 123 cases (97%), and the ability question was filled out in only 92 cases (75%).

The box plots of the OSATS scores corresponding with the ability to operate autonomously are presented in figure 2. The median OSATS score that corresponded with the supervisor's opinion that the resident was able to perform the procedure autonomously was 28 (range 20-30). This corresponds with 92% of the maximum score, taking the possible range from 6 to 30 into account.

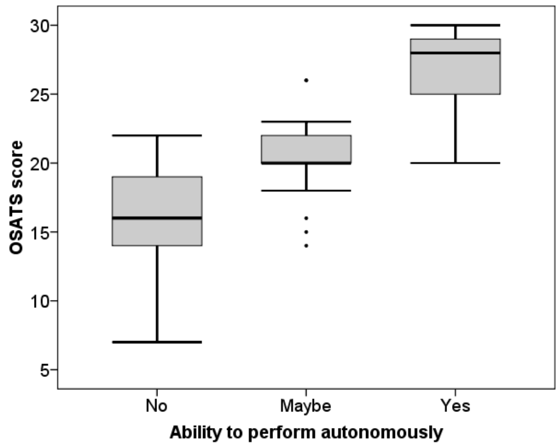


Figure 2. OSATS score against the ability to perform autonomously.

Box plots of the OSATS scores as rated by supervisors against their answer on the question whether to resident is able to perform the procedure autonomously. The lower and upper lining of the boxes represent the 25th and 75th percentile, the black line is the box is the median. The number of procedures rated with no, yes and maybe were 58, 36 and 28 resp.

The ICC of the total OSATS score of residents and supervisors was 0.78 (95% CI 0.70-0.84), indicating substantial agreement. Additional analysis was carried out on the cases in which residents overrated versus underrated their performance (defined as a resident's OSATS score of ≥ 3 versus ≤ 3 points compared to the supervisor's rating). According to this definition, overrating was present in 15% (n=18), agreement in 53% (n=66), and underrating in 32% (n=39) of the procedures (Figure 3). In seven of 18 cases (38%) in which the resident overrated his/her performance, the resident was in the last two years of residency training (PGY5 or PGY6). Next, supervisors and residents agreed in 64 out of 91 (70%) cases regarding the question of whether

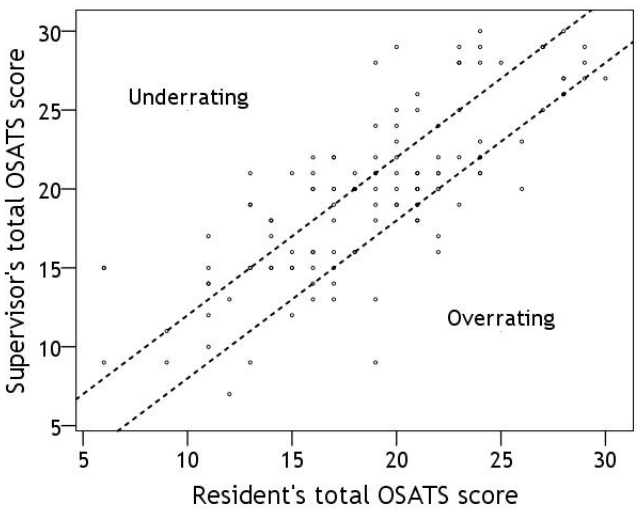


Figure 3. Supervisor's OSATS score plotted against Resident's OSATS score. Scatter plot. The area between the dotted lines represents agreement (a difference of 2 or less points). Under- and overrating of the resident is defined as a self-perceived OSATS score of at least 3 points lower resp. higher than the supervisor's score.

a resident was able to perform the procedure autonomously. No absolute disagreement occurred (Table 1).

Table 1. Agreement resident and supervisor.

Agreement Supervisor and Resident regarding ability to perform procedure autonomously					
		Ability according to resident			
Count		No	Maybe	Yes	Total
Ability according to supervisor	No	38	12	0	50
	Maybe	5	9	5	19
	Yes	0	5	17	22
	Total	43	26	22	91

The response rates to the questionnaire were 96% and 100% respectively: 23 residents and 15 supervisors returned the questionnaire. One resident did not respond due to inexperience with the OSATS because of having just started the residency training. The results are presented in table 2. Regarding the ideal frequency of administration of an OSATS, the majority of the residents (17 out of 23) answered that this was after each procedure they had performed as the primary surgeon.

Table 2. Results survey.

OSATS ...		Median score residents	Median score supervisors
is a valid instrument	(1= strongly disagree; 5=strongly agree)	3	3
is (subjective/objective)	(1=subjective; 5=objective)	2-3	3
should be used for assessment	(1=strongly disagree; 5=strongly agree)	4	4
helps acquiring surgical skills	(1=strongly disagree; 5=strongly agree)	3-4	3-4
leads to irrelevant paperwork	(1=strongly disagree; 5=strongly agree)	2-3	2

DISCUSSION

At 92% of the maximum OSATS score (28/30 points), a supervisor found a resident Ob/Gyn is able to perform a procedure autonomously, irrespective the kind of procedure. This is a rather high score, especially in comparison with a randomized clinical trial (RCT) that described a cut-off point for a certain level of proficiency.[Bijen et al., 2009] The authors used 75% of the maximum OSATS score (28/35 points) to select surgeons as proficient to perform a total laparoscopic hysterectomy in the case of endometrial cancer. The high score measured in this study might be due to the supervisor's striving for perfection and his/her hesitancy to authorize the resident to perform the surgery autonomously. Furthermore, it may also be the result of the desire to add something to a resident's learning process, and not wanting to be redundant as a teacher. Also, an assessing supervisor in our study may have interpreted the maximum score as the level required at the end of residency, while an assessor during the clinical trial[Bijen et al., 2009] may have taken the maximum score as absolute perfection. Assessing an individual with a certain frame of reference in mind diminishes the objectivity the assessment instrument.

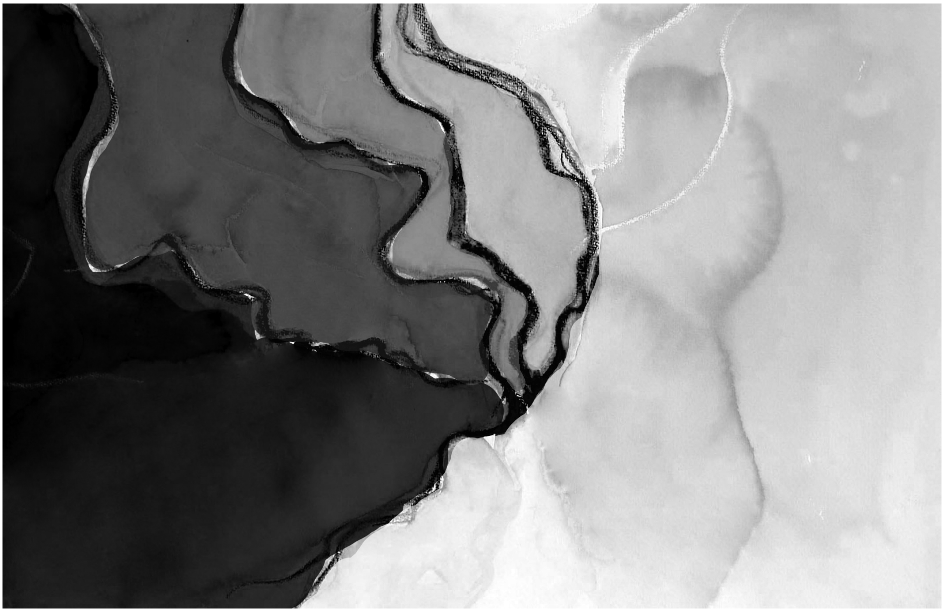
Substantial agreement is present between residents' and supervisors' OSATS score. This is in accordance with the inter-observer reliability that is found in laboratory settings,[Martin et al., 1997] and endorses the quality of OSATS as intraoperative assessment tool. Residents, however, cannot be regarded as equivalent to observers since they performed a self-assessment. Regarding self-assessment, the study of Mandel et al. found good reliability comparing the rating of trained faculty observers with residents' self-assessment of surgical skills on bench models using the OSATS.[Mandel et al., 2005] However, a recent review raises questions about the abilities of health professionals to generate accurate judgments of their own performance.[Eva et al., 2005] Even in concrete areas such as technical knowledge and ability, the self-assessment was found to be inaccurate.[Ginsburg et al., 2000; Gordon, 1992; Hodges et al., 2001] An issue of greater concern is that those who perform worst on external assessment may also overrate their performance on self-assessment.[Davis et al., 2006; Hodges et al., 2001; Kruger et al., 2006; Lynn et al., 2006] Fortunately, we showed that only a minority (15%) overrated themselves. Nonetheless, overconfidence is dangerous, especially when combined with suboptimal performance at the end of residency training. Notwithstanding, self-assessment has assumed increasing importance, though external assessment will always play an essential role during the process of acquiring certain skills.[Mandel et al., 2005]

Official assessments generate paperwork. However, neither residents nor supervisors agreed with the statement that OSATS leads to irrelevant paperwork. Instead, they agreed that it should be used for assessment and evaluation, and that it helps to improve residents' surgical skills. Furthermore, residents state that they want to be assessed after every surgical procedure they performed as primary surgeon. In daily practice, however, they only request this during a minority of the procedures performed. Probably, their answers express socially desirable behaviour, or there may be barriers present that discourage residents from asking for an assessment with an OSATS. A possible explanation is that practical impediments (e.g. insufficient time) hamper the frequency of administration of the OSATS. On the other hand, it is questionable whether all procedures need to be assessed. In our opinion, regular assessments distributed over time are sufficiently able to show a resident's skills level as well as the expected progression in skills and in the performance of the surgery. Worthwhile situation-specific feedback and advice need to be given during every surgical procedure. Strikingly, none of the participants judged the OSATS to be a very objective assessment tool. This is in contradiction with the objectivity the OSATS stand for. Obviously, no judgment from one person about another will be free from subjectivity. Though, in our opinion, this finding suggests that someone's surgical performance cannot automatically be derived from an OSATS score.

This study addresses a very difficult area of surgical skills evaluation, and was conducted under regular conditions mirroring daily practice in a residency program. Following such a design, real-life influences were allowed to colour the study results. Individual variation will have been present among residents, supervisors and procedures. It is unlikely that all residents had the same level of motivation to participate in the study. This was illustrated by the large range of number of OSATS each resident had collected. Additionally, variation will have been present in the extent to which supervisors allowed residents to independently perform a surgical procedure and in their method of assessment. Also, the procedures assessed will have varied widely with respect to difficulty and the risk of complication. However, the OSATS will be implemented in the actual clinical situation and not under predefined study conditions. This expresses the strength and the value of this study.

Training in all specialties is evolving and moving towards more competency-based outcome measurements rather than solely based on the length of training. This is a positive development. However, we should avoid an indiscriminate implementation of instruments such as the OSATS, especially, with respect to drawing consequences to certain scores like authorization. Authorization is a more complex, multifactor process. During this process, the importance of an OSATS is limited. Other competencies also have to be taken into account, such as knowing when to operate and when not, recognizing someone's own abilities and inabilities, asking for help when needed, and being open to suggestions from colleagues. Therefore, all these competencies should be evaluated prior to authorization. Finally, acquiring surgical proficiency is an ever-continuing process that does not end with the completion of residency training.

CHAPTER 10
ARE MINIMALLY INVASIVE PROCEDURES HARDER TO ACQUIRE THAN CONVENTIONAL SURGICAL PROCEDURES?



Ellen Hiemstra
Wendela Kolkman
Saskia le Cessie
Frank Willem Jansen

Adapted from Gyn Obstet Invest. 2011; 71: 268-73

INTRODUCTION

The psychomotor skilfulness of a surgeon reveals to be one of the most important factors influencing the outcome of a surgical procedure, combined with his or her cognition and personality.[Najmaldin, 2007] Nevertheless, operating room (OR) experience is harder to gain for residents these days. This is mainly due to reduced working hours, financial and ethical constraints.[Hammond & Karthigasu, 2006] Consequently, in order to ensure safe and high quality treatment, the skills training needs to be optimized by means of structured programs and by the implementation of appropriate feedback and assessment during residency to objectively evaluate surgical skills.[Aggarwal et al., 2004]

Recently, an assessment tool has been implemented to objectify surgical skills: the Objective Structured Assessment of Surgical Skills (OSATS). This tool was originally designed to measure technical surgical performance in skills laboratories, and its validity, reliability and feasibility have been established in these settings.[Martin et al., 1997] Additionally, Aggarwal et al. have validated the general global rating scale of the OSATS for real laparoscopic surgery.[Aggarwal et al., 2008] For their validation they used video material recorded by the laparoscopic stalk during laparoscopic cholecystectomies. Furthermore, the OSATS have been validated for use during actual observation during surgery by our study group.[Hiemstra et al., 2011]

The advent of minimally invasive surgery (MIS) has even further increased the interest in skills training and assessment, because it is considered to be a more demanding technique for surgeons than the conventional (open abdominal and /or vaginal) surgery.[Feldman et al., 2004b] Arguments for the complexity of the MIS technique are the long surgical instruments that are required with reduced haptic feedback and fewer degrees of freedom, the altered depth perception resulting from 2D imaging and the necessity to adapt to the fulcrum effect.[Gallagher et al., 1998; Perkins et al., 2002] Therefore, it has often been concluded that MIS requires a longer learning curve than conventional surgery.[Moore et al., 1995; Perkins et al., 2002; Purkayastha et al., 2004] However, this assumption is based on arguments rather than objective measurements.

Although insight in the process of acquiring surgical skills is highly important, data on objective measurements on the surgical learning process during residency are scarce. Therefore, this study was conducted to gain a better insight in the residents' learning process of technical operative skills. Specifically, we tried to find support for the assumption that residents experience more difficulty to acquire MIS procedures than conventional surgical procedures by means of OSATS-based learning curves.

MATERIALS AND METHODS

An observational cohort study was conducted at our university teaching hospital, the Leiden University Medical Centre (LUMC). In general, residents in Obstetrics and Gynaecology (Ob/Gyn) spend three years of their six-years' residency program in a university teaching hospital to be trained in a variety of subspecialties, like reproductive health care, perinatology and oncology. A three months clinical rotation is spent on gynaecological surgery, generally attended during the fourth postgraduate year (PGY 4). They had gained some prior surgical experience during the first year of training which is spent in the general Ob/Gyn practice in a non-university teaching hospital. This experience was limited to urgent surgery on call, and almost no elective surgery. The second and third year they are mainly trained in obstetrical skills in a university teaching hospital. During the clinical rotation 'gynaecological surgery' during the fourth year of residency, they are scheduled to perform surgery in the OR for four or five days a week. For the numbers and specific type of procedures, they depend on the normal throughput of patients scheduled on the operation program. Parallel to their increase in experience, they gradually perform each procedure more autonomously, depending on their level of performance and patient characteristics. As their increase in responsibility was merely based on the supervisor's general opinion, rather than on more objectively defined measures, this can be taken as training according to the conventional apprenticeship model.

The general global rating scale of the Objective Structured Assessment of Technical Skills (OSATS) was included in the intraoperative assessment of surgical performance. The assessment form was adapted from Martin et al.[Martin et al., 1997] Six domains of surgical technical competence are scored on a 1 to 5 Likert scale, with explicit descriptions at point 1,3 and 5. Originally, the OSATS had been designed to rate surgical skills in skills laboratories. In the Netherlands, this evaluation method has largely been implemented for assessment purposes in the OR.

During a 27-months investigation period, each resident that consecutively started the clinical rotation in gynaecological surgery was asked to participate in the study. They were instructed to ask the consultant, who was scheduled as supervisor in the operation room, for an assessment with an OSATS after every procedure that they performed as a primary surgeon during this period. Procedures during which a resident independently performed some important steps were included for assessment as well. Supervisors were instructed to fill out the OSATS form by rating the performance on each domain, irrespective of the resident's training level.

Data were analysed using SPSS for Windows (SPSS version 16.0 SPSS Inc., Chicago, IL). The total score of each OSATS was calculated by adding up the score of the six domains (at minimum 6 and maximally 30 points). Learning curves were drawn by plotting resident's OSATS scores against his/her procedure-specific experience, in which the experience was quantified by the surgical caseload (one number was added to the caseload for each consecutive procedure that had been rated with an OSATS). The curve for the mean OSATS score per caseload was plotted to approximate the general learning curves for MIS and conventional surgery. To study the relation between OSATS score and caseload for the different surgical techniques, linear mixed models (LMM) were used. These models were fitted as a random coefficients model - a random slope and a random intercept - for resident and a fixed effect for the type of procedure. P-values <.05 were considered statistically significant. Ninety-five per cent confidence intervals (95%CI) were calculated.

RESULTS

All nine residents who attended their three months clinical rotation in gynaecological surgery during the investigated period agreed to participate in the study. Although no exact data of all participants was available, their prior experience was limited to about ten autonomously performed caesarean section, and some diagnostic laparoscopies or the removal of an ectopic pregnancy. In general, they had not performed elective surgery yet, and therefore no experience with vaginal or hysteroscopic procedures. Obviously, inter-individual variations would have been present.

A total of 319 surgical procedures were assessed; 129 OSATS for MIS and 190 OSATS for conventional surgery, 40 and 60% respectively (Table 1). Regarding MIS, 98 laparoscopic and 31 hysteroscopic procedures were assessed. The majority of the laparoscopic procedures included removal of an adnex, cystectomy, diagnostic laparoscopy and tubal sterilization. The hysteroscopic procedures were diagnostic and therapeutic, in which the latter mainly concerned resections of polyps and myomas type 0. The conventional procedures were either performed using an abdominal (n=125), or a vaginal approach (n=65). Conventional procedures were, next to caesarean sections, mainly abdominal and vaginal hysterectomies. An individual resident obtained a median of 40 assessed procedures (range 12-60), of which 13 procedures were minimally invasive (range 2-26).

10

Table 1. Characteristics of learning curves MIS and conventional surgery.

Category of procedures	Number of assessed procedures	Learning curve characteristic	
		Intercept (95%CI)	Slope (95%CI)
MIS procedures	**129**	**17.2 (15.3-19.2)**	**1.77 (1.19-2.35)**
Laparoscopic procedures	98	18.9 (16.7-21.2)	1.40 (0.16-2.63)
Hysteroscopic procedures	31	12.9 (9.5-16.3)	2.69 (1.45-3.94)
Conventional open procedures	**190**	**21.5 (19.6-23.3)**	**0.75 (0.15-1.35)**
Abdominal approach	125	21.5 (19.0-24.0)	0.69 (0.18-1.20)
Vaginal approach	65	21.7 (19.1-24.3)	0.47 (-0.05-1.00)

The supervising consultants were 21 gynaecologists. Some of them were obstetricians who only supervised obstetric procedures like caesarean sections, while others were experts in minimally invasive surgery who also supervised conventional surgical procedures.

The total OSATS scores were plotted against a resident's procedure-specific experience for the first ten procedures. Additionally, the average MIS OSATS score and conventional OSATS score were calculated for each caseload (1 to10). The resulting two curves can be interpreted as an approximation of the general learning curves for both surgical techniques (Figure 1).

The average OSATS score plotted against procedure-specific experience. MIS = Minimally invasive surgery (laparoscopy/hysteroscopy); conventional surgery = open abdominal/vaginal procedures.

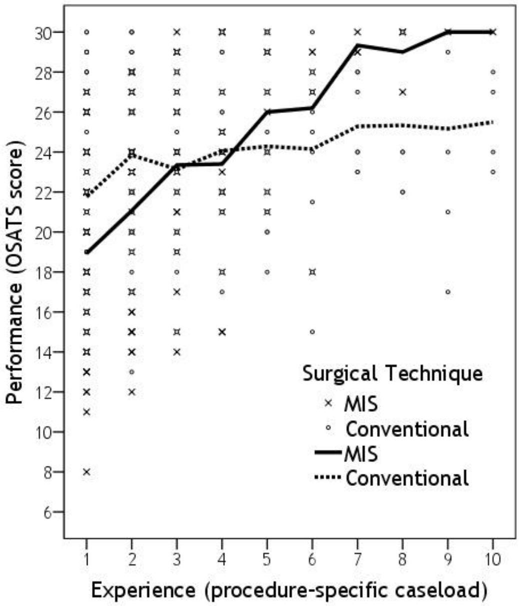


Figure 1. Learning curves for MIS and conventional surgery.

LMM analysis revealed that the slopes of the learning curves, i.e. the average increase in OSATS score for each consecutively assessed procedure, differed significantly for MIS versus conventional surgery, 1.77 versus 0.75 points per procedure (95%CI: 1.19-2.35 versus 0.15-1.35, $p<.01$). Table 1 shows the slopes as well as the intercepts resulting from the LMM analyses.

DISCUSSION

Residents in Ob/Gyn progress at least as fast along the learning curve for MIS as along the curve for conventional procedures during an intensive three-months clinical rotation in gynaecological surgery. This finding is in contrast with the often heard, but never scientifically supported, concern that surgeons have to proceed along a longer learning curve to acquire these MIS skills.

Concerns about how to acquire these complex skills should be considered in the context of the explosive growth of the MIS technique. In a relatively short time this approach has evolved to be the ‘gold standard’ for many disorders, like ectopic pregnancies and benign ovarian tumours.[Clasen et al., 1997; Medeiros et al., 2008] However, after the initial reports of success of MIS, doubts surfaced regarding its safety.[Aggarwal et al., 2004] A factor possibly contributing to these doubts is that the surgeons, although experienced in conventional surgery, often had to acquire the MIS skills in an autodidactic way. At these times, neither structured training programs, nor simulators were available to train their skills. Under those circumstances, it was probably hard to transfer the skills which they had just acquired themselves to the residents in training. This factor certainly has contributed to the slow implementation of MIS techniques, as observed in the Netherlands.[Kolkman et al., 2006]

For this study, we choose to classify both laparoscopy and hysteroscopy as MIS. Even though these surgical techniques are different, they require common psychomotor skills, like manipulation of long surgical instruments while looking to a video screen and interpretation of a three-dimensional operation field from a two-dimensional display.

For interpreting the faster advancing of residents along the MIS learning curve than along the curve for conventional procedures four factors need to be taken into account: the resident seniority, a structured curriculum, their audio-visual dexterity, and the mentor proficiency. In the first place, they all were PGY 4 residents, and already had gained some operative experience in their preceding three years of residency. On call, residents were equally exposed to conventional open surgery, e.g. caesarean sections, as to MIS, like a diagnostic laparoscopy or laparoscopy for an ectopic pregnancy. This implies that conventional surgery is as 'conventional' to them as MIS. Secondly, they all had been exposed to simulator training focusing on MIS, because they all had attended the uniform, mandatory Dutch basic skills course during the first two years of residency.[Hiemstra E et al., 2008] During that course, they attended theoretical sessions and received hands-on training on validated endoscopic trainers. Technical skills acquired on validated simulation devices have proven to be transferable to the OR setting. [Anastakis et al., 1999; Hyltander et al., 2002] In the third place, this younger generation of residents has, at average, experienced an earlier introduction to computers and other audio-visual devices. This will thin the technical interface between surgeons and screen-mediated medical applications, like hysteroscopy and laparoscopy.[Rosser, Jr. et al., 2007] This may have led to an easier acquisition of MIS skills for current generation of residents. Finally, the teachers of the MIS procedures will have played an important role. Currently, they may be better able to transfer their skills to the next generation compared to a decade ago when they were still progressing along their own learning curve.

The diversity of the study population is inherently to the clinical research that was performed. The nine participants considerably varied with respect to the number of conventional and laparoscopic procedures they had performed prior to inclusion in the study. Furthermore, their manual dexterity, their eagerness to acquire surgical skills varied and to collect OSATS forms was not consistent. This can be taken as a limitation of the study. However, as in daily practice all these influences were allowed to colour the results, this study generates unique and actual information about clinical practice when compared to data collected in laboratory settings.

However, it has to be considered that the residents' palette of MIS surgery mainly consisted of adnexectomies and cystectomies, while the majority of the conventional procedures were caesarean sections and hysterectomies, performed either abdominally or vaginally. Obviously, the two categories of procedures are not equivalent. Intuitively, these MIS procedures are less complex than the conventional procedures. To our knowledge, no system is available to compare the complexity of all these procedures. However, the European Society of Gynaecological Endoscopy (ESGE) classified laparoscopic procedures, and a cystectomy and an adnexectomy are less complex procedures than the laparoscopic hysterectomy (respectively level 2 and 3).[ESGE 2009] Nevertheless, it is promising to observe that the basic MIS procedures are not difficult to learn within a three-months clinical rotation. Hopefully, this will result in a speedier implementation of the MIS technique by the next generation of gynaecologists.

Assessment of surgical skills is a very important and contemporary concern for all teachers in surgical professions.[Chen et al., 2010; Sweet et al., 2010] Although OSATS have been designed to assess skills objectively, an assessment by an individual person is by definition not free of subjectivity. Therefore, the objectivity of OSATS to measure learning curves may be criticized. Though, in absence of a gold standard to assess OR performance, it may be the best tool currently available. At least, it has proven to be superior to a task specific checklist and a pass/fail judgement,[Martin et al., 1997] and it surely is less subjective than a general assessment at the end of a rotation on a recall basis. Furthermore, other measures to monitor the learning process like the duration of the procedure and the complication rate do not seem that useful during residency because these largely depend on the supervisor. In fact, it is questionable whether these measures reflect a surgeon's skilfulness at all, being influenced by the selection of the patients.

The OSATS-based learning curves of PGY 4 residents during three-months clinical rotations form an important first step to gain insight in residents' learning process in the OR setting. The results indicate that basic MIS procedures are not harder to acquire during residency than conventional surgical procedures. Moreover, as the current residency program rather well facilitates the acquiring of basic MIS procedures, residents are provided with a solid foundation to progress to the more advanced MIS procedures after residency.

CHAPTER 11
GENERAL DISCUSSION

Frank Willem Jansen
Ellen Hiemstra

Adapted from BJOG 2011, accepted

Minimally invasive surgery (MIS) presents new technical challenges for the surgeon. In fact, the visualisation of the operative field is completely different and the surgical instruments have to be handled remotely. The need to benefit maximally from the learning moments in the OR whilst optimizing patient safety and the particular skills required for MIS have led to the development of simulators. Despite the compelling arguments to support the widespread use of these simulators as a core and mandatory part of MIS training, the implementation has been lagging behind.

In this thesis, a scientific basis is set for the organisation of MIS skills training. Hopefully, this will help to guide the demands of occupational groups of medical specialists and government for a more uniform and better implemented training regimen. Nowadays, this is important with the increasing pressure to assess the quality of health care. The ultimate goal is to enhance patient safety. When setting up a skills training program, the curriculum needs to be carefully considered in terms of specific and measurable learning outcomes. Validated training and assessment tools should be employed where available. Not only can simulation develop, consolidate and evaluate surgical skills, but it could also be used to identify a trainee's deficiencies and qualities. The latter will allow for the development of personalized training regiments. Finally, skills learned outside of the OR need to be integrated into the live situation. Findings and considerations related to the optimal organisation of skills training and assessment are discussed below.

11

Curriculum development

In the absence of a structured curriculum, a well-equipped skills laboratory does not guarantee success. Up until now, standards or guidelines of how the MIS skills training should be organized are lacking. Therefore, we developed an international and consensus based set of quality criteria for a skills laboratory. According to worldwide experts in MIS, the most important factor to motivate surgical trainees is mandatory training supervised by laparoscopic experts. **(chapter 7)** Training facilities remain unutilized if practice of MIS skills is considered voluntary.[Chang et al., 2007] It is not only essential that the training is mandatory, the performance of laparoscopic surgery in the OR should only be sanctioned once trainees have achieved a predefined skill level.[Stefanidis & Heniford, 2009a]

The most important criteria regarding curriculum development are the presence of a structured skills curriculum, dedicated time for skills training and a yearly evaluation of the progress and maintenance of laparoscopic skills of the resident. **(chapter 7)** In addition, the retention of acquired skills needs to be monitored. This is important, because obtaining a diploma for basic laparoscopic skills once, does not guarantee that the skills remain over time. Trainees Obstetrics and Gynaecology often encounter periods of non-exposure to the OR, given the rotation of residents around the variety of disciplines and hospitals. We showed that tissue handling skills diminish slightly in absence of training or patient exposure of one year, because a deterioration of time and precision was observed to perform a task that mainly required those skills. **(chapter 6)** Therefore, an annual evaluation seems appropriate. In accordance with that, a recent study confirmed that laparoscopic skills deteriorate between 6 and 18 months without training.[Maagaard et al., 2011] This additionally supports a re-examination after one year.

Appraising the organisation of MIS skills training in the Netherlands, it can be determined that the demands of the Dutch Health Inspectorate for uniformity in training and a predefined skills level are met by the mandatory nationwide basic surgical skills course. **(chapter 2)** However, the additional implementation of skills training in the curriculum is left to the individual clusters of teaching hospitals. Regarding the latter, the ranked list of quality criteria for MIS skills laboratories **(chapter 7)** can be used to verify the quality of an existing laboratory. From there, the focus for new developments can be chosen to upgrade the quality of the lab at specific levels. In general, the adage "See one, do one, teach one" needs to be changed to "See one, do multiple in a skills lab, do one for real" [Bashankaev et al., 2011]

Type of trainer model

Animal models most closely approximate operating on a live patient in terms of being the only models that can effectively simulate a bleeding and complications. However, they are expensive and are associated with infectious, moral and ethical concerns. [Hammoud et al., 2008] Furthermore, the anatomy of a human female genital tract is not equivalent to a pig's model. Regarding inanimate models, the two available trainer models are box and VR trainers. Although evidence is convincing that both models improve psychomotor surgical skills, they have different characteristics. For one, box trainers are much cheaper than VR trainers. During box training, surgeons can familiarize themselves with real laparoscopic instruments and natural haptic feedback (in terms of the feel of the instruments on the tissue surfaces and the pressure of opening and closing the instruments) is preserved. Haptic feedback is especially important during the early phase of psychomotor skill acquisition.[van der Meijden et al., 2009] The absence of haptic feedback during VR training is a disadvantage when trying to replicate traditional MIS, but it may be a truer representation of robotically assisted MIS in which haptic feedback of the tissue is lacking.[Hammoud et al., 2008] In VR trainers, entire procedures can be trained. A recent study showed that training laparoscopic salpingectomy in VR leads to an improvement of skills level during the real procedure.[Larsen et al., 2009] The authors suggested that procedure specific training improves cognitive skills in addition to psychomotor skills. Moreover, most VR trainers give instant feedback on performance allowing solitary training and personal assessment. During box training, this is only possible if the box is equipped with a tracking device that generates instant feedback. However, verbal expert feedback turned out to be superior to computer generated feedback in terms of economy of movement, especially with respect to retention of skills. [Porte et al., 2007] In addition to provision of individualized feedback, expert feedback allows for the opportunity to exchange tips and tricks for daily clinical practice. Finally, portable box trainers with fixed video cameras can also provide the opportunity for practice 'at home'. In summary, the superiority of one of both simulators has been disputed in multiple studies, often with poorly comparable training setups with varying outcomes. We found that box training models are superior to traditional VR systems for an exercise in which tissue handling is important. However, additional kinematic interaction between instruments and objects can be a promising surrogate for haptic feedback in VR systems. **(chapter 3)**

Based on their different characteristics and varying advantages, we conclude that it is ideal to use box and VR trainer in tandem. Future evidence should be sought to ascertain how they should ideally be combined within training programs. However, it has to be considered that the box trainer is the authority based standard, derived from the higher priority that worldwide experts in MIS training gave a to the presence of a box trainer than to the presence of a VR trainer. **(chapter 7)**

Exercises

It is of the utmost importance that the construct validity of the exercises used, is confirmed. Preferably, an expert standard has been established to allow goal oriented training, and thereby fuelling the motivation of participants.[Stefanidis & Heniford, 2009a] When making a choice of the available validated exercises, the learning objective of different exercises needs to be taken into consideration. Some tasks focus, for example, predominantly on hand-eye coordination, while tissue handling is more important in others. Unfortunately, the effectiveness of different exercises has not been compared yet, in spite of some reports of participants' preferences. Also, the transferability of exercises to the real laparoscopic situation is relevant for selecting exercises. Cutting, suturing and knot tying are directly transferable to a real laparoscopic procedure. For that reason we showed the validity of the clinically relevant knot tying task. (**chapter 5**) In addition, future research should be undertaken to categorize the various training exercises and examine their effectiveness.

Metrics

A simple and feasible measure of assessment is the time needed to complete a task. Although in general time is able to discriminate between surgeons of different skill levels, a measure of precision should be added from a clinical point of view. Additionally, it was revealed that it takes longer to achieve precision than speed. [Smith et al., 2002] Therefore, Kolkman et al developed a composite score rewarding speed and precision for the five validated exercises used in this thesis.[Kolkman et al., 2008] Despite the superiority of time as a measure of assessment, trainees benefit from feedback on performance in the form of motion analysis parameters. [Stefanidis et al., 2009b] Also, motion analysis parameters discriminate between surgeons of various skill levels during real laparoscopic procedures.[Smith et al., 2002] We proved the construct validity has for time, path length, motion in depth, and motion smoothness of the laparoscopic suturing task using a box trainer. The addition of economy of movement as a measure of assessment to time to complete the task has the potential to refine acquisition of skills.**(chapter 5)**

Time taken to complete the task, precision and economy of movement parameters may suffice for many exercises. However, these outcome measures are less appropriate in the case of a task during which tissue handling is predominantly required. Force imparted by the operator on the tissue is likely to be of greater importance and attempts have been made to measure forces used during the performance of certain exercises.[Horeman et al., 2010] Research needs to be done on how force application should be integrated into laparoscopic skills training.

Selecting trainees

Ex vivo training and assessment has the potential to contribute to the selection of appropriate candidates for surgical residency positions. Stefanidis et al. were able to predict the rapidity of skills acquisition based on simple psychomotor tests.[Stefanidis et al., 2006a] A recent study also found a correlation between innate psychomotor and visuospatial abilities and skill level at the end of a training session.[Van Herzeele et al., 2010] Certainly, identification of residents' particular strengths and weaknesses may allow for a more tailored, individualized approach to training, but whether tests of baseline and rudimentary performance on laparoscopic simulators can find a valid role in candidate selection remains a subject for debate. No doubt, skilfulness is a prerequisite for a surgeon. However, there is no guarantee that the medical student with the best results in a psychomotor test, like a finger taps test, will also become the best surgeon during residency. In fact, master surgeons recognize cognitive factors and personality (decision making ability, insight, team spirit, and emotional stability) as being of equal importance for selection [Cuschieri et al., 2001] Currently, differentiating between better and less skilled trainees seems more important for tailoring training to individual needs, rather than for selection purposes.

Assessment in the OR

The organization of skills training outside of the OR, including its implementation into the residency curriculum, is an essential first step to warrant patient safety during MIS. After basic surgical skills have been acquired outside of the OR, residents should be trained, assessed and reassessed in the OR in order to achieve transparent skilfulness. This parallels the development that all medical specialties move towards more competency-based outcome measures rather than being solely based on the length of training or the number of procedures performed. In an attempt to overcome the subjectivity on which surgical skills assessment was traditionally based, the general global rating scale of the OSATS has been developed for testing in laboratory settings. Two findings indicate the construct validity of using the general global rating scale of the OSATS for intraoperative assessment. Firstly, surgeons who had performed more than 100 laparoscopic cholecystectomies were rated with a higher score than surgeons who had performed less than 10, based on video assessment.[Aggarwal et al., 2008] Secondly, the OSATS score, obtained by direct intraoperative observations, raises with increasing experience of individual residents **(chapter 8)**. Furthermore, the reliability of assessment with OSATS has been proven in laboratory settings.[Martin et al., 1997] In accordance, we confirmed substantial agreement between resident and supervisor if an OSATS is used for intraoperative assessment **(chapter 9)**. The ideal proof would be to schedule two independent supervisors during a series of procedures. However, this is a very costly method conflicting with the strive for efficient health care. Therefore, the present evidence for reliability is the best available. The large scale implementation of OSATS for intraoperative assessment is permitted with respect to its validity and reliability, although it already took place prior to manifestation of the results mentioned above.

Additionally, practical decisions have to be made about the intraoperative use of OSATS. The timing and frequency of assessment remains controversial [Pandey et al., 2006] The majority of questioned residents state they prefer every procedure to be assessed. (**chapter 9**) Although it is well established that feedback is indispensible during learning,[Mahmood et al., 2004] it remains questionable if residents who are assessed with OSATS will become better surgeons than those who are not. Furthermore, it is questionable whether consequences should be drawn from a residents OSATS score and what these consequences should be. We have made a first step to benchmark the OSATS, by establishment of the score at which a resident can perform a procedure autonomously (**chapter 9**). However, a thoughtless implementation of such assessment tools for authorization should be avoided. In the first place, authorization is a more complex and multifactor process which depends on more aspects than technical skills only, e.g. knowledge, decision making (before, during and after the operation), communication skills, and leadership skills.[Moorthy et al., 2003] In the second place, the objectivity of the OSATS is limited, as appeared from a survey among users (**chapter 9**) and the large range of median OSATS scores given by the supervisors (**chapter 8**).

An advantage of the fact that OSATS has proven to be valuable is that it fuels other clinical research regarding technical skills performed during actual surgery. For example, we can be reassured with the finding that basic MIS procedures do not seem harder to perform during residency than conventional surgical procedures.(**chapter 10**) This may have resulted from the incorporation of structured MIS training programs in residency. It is likely that the current generation of residents familiarizes quicker with new technologies due to an earlier introduction to computers and other audiovisual devices. This is illustrated by the findings of Rosser et al. that gaming surgeons operate quicker and with less errors.[Rosser, Jr. et al., 2007]

This thesis has focused on training and assessment of psychomotor technical surgical skills. Obviously, in an era of rapid technological innovations, a surgeon should also be well informed about instrumentation and OR set up (e.g. integrated operating rooms). However, the time has passed when medical education could be planned with a focus solely on the latest aspects of medical diagnosis and (surgical) treatment. The Royal College of Physicians and Surgeons of Canada (RCPSC) has developed the CanMEDs competency framework for the education of medical professionals. Seven roles have been defined that a physician should fulfil to meet the health care needs of the patients, communities and societies they serve. These roles are medical expert (central role), communicator, collaborator, manager, health advocate, scholar and professional.[Frank et al., 2007] As a result, residency programs are now restructured with more competency based training and outcome measures. Residents that excel in one area of competence but lack in other competencies may need to shift their priorities in their training curriculum. [Schijven et al., 2011] Examinations are indispensible in that process and may function to tailor the curriculum further to each individual.

Finally, acquiring surgical proficiency is an ever continuing process that does not end with completing residency training. The requirements in training and assessments for residents will also have implications for use in revalidating more experienced surgeons. In parallel, surgeons should also be prevented from implementing new technologies (e.g. laparoscopic endoscopic single site surgery and robotically assisted surgery) without proper training and transparent proficiency. Unfortunately, consensus regarding the latter still has to be achieved.

CHAPTER 12
CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS AND RECOMMENDATIONS

From literature search and based on the results of this thesis the following is recommended:

Training facilities outside of the OR

- Acquiring the manual dexterity for successful MIS has a marked learning curve and should be done before starting to operate on an actual patient.
- For training basic laparoscopic skills a physical box trainer is an appropriate and cheap solution. The skills acquired on that trainer are transferable to the OR setting and tissue handling can be trained using real laparoscopic instruments. The camera setup in the box trainer can be either a fixed or a navigated system.
- Path length, motion in depth and motion smoothness are motion analysis parameters that have been validated for the clinically relevant knot tying task. These parameters can be retrieved in a box trainer equipped with a tracking device. The use of an expert standard based on motion analysis fuels motivation, because it serves a training goal. Furthermore, it potentiates the refinement of psychomotor skillfulness. Finally, assessment using forces applied to the tissue is potentially worthwhile. This topic requires further exploration.
- Skills training should be embedded in a curriculum in which residents should be obliged to practice until a predefined level of skills has been met, prior to performing surgery on a living patient. A yearly retention of skills measurement should be included.

The minimally requested items for a MIS skills laboratory are listed in table 1.

Table 1. Minimally requested items for a MIS skills laboratory and curriculum.

availability financial resources
presence of a box trainer
presence of a curriculum director (a laparoscopic expert)
training sessions are supervised by a laparoscopic expert
training is mandatory
residents are not allowed to perform surgery if predefined skills level is not reached
presence of a structured skills curriculum
time is dedicated for skills training
maintenance of training is embedded in the curriculum
progress in laparoscopic skills is incorporated in yearly evaluation of resident

Assessment using OSATS

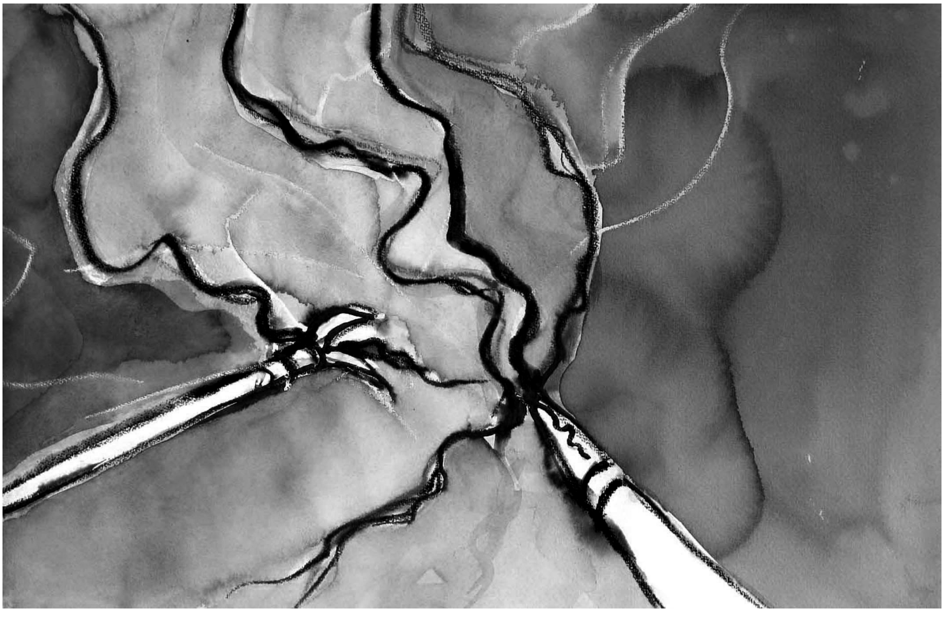
- The OSATS is a valid, reliable and feasible instrument for the evaluation of intraoperative surgical skills for both MIS and conventional surgery. However, concerns have arisen about its objectivity to measure skillfulness.
- The total OSATS score correlates with the finding that a resident can perform a procedure autonomously. Though, OSATS should not be the only tool used for authorization.
- After a proper preparation in a skills laboratory, MIS procedures are not harder to acquire than conventional surgical procedures
- Although many residents claim they would like to be assessed with an OSATS after every procedure, this frequency is not necessary to express a learning curve or to indicate the average technical skillfulness of an individual resident.
- The steps required for the implementation of the OSATS in the residency curriculum are listed in table 2 and is adapted from the Dutch Journal for Medical Education[Hiemstra et al., 2010].
- Assessment with OSATS meets the societal demand for transparency in medical care. Structured feedback has the potential to aid the proficiency gaining curve. However, whether OSATS will lead to technically better surgeons is uncertain and might be subject of future research.

Table 2. Stepwise implementation of the OSATS.

Step 1	Decide with (cluster of) teaching hospital(s) which procedures should be assessed with an OSATS
Step 2	Inform, instruct and motivate resident and supervisor. Only instruction needed is "assess irrespective the training level".
Step 3	Guarantee access to the assessment forms (e.g. electronically)
Step 4	Benefit from the possibilities incorporated in assessment with OSATS

12

CHAPTER 13
SUMMARY / SAMENVATTING



SUMMARY

The background of the studies presented in this thesis is given in **chapter 1**. Minimally invasive surgery (MIS) was developed as an alternative for large incisions (e.g. laparotomy) in order to create as little tissue damage as possible. As a result, the MIS technique has advantages for patient recovery and cosmetics. However, the surgical technical skills needed to perform this form of surgery safely and efficiently are more challenging. The operation field is not directly visualised, but by means of a camera image projected on a 2D screen. Next, the gloved surgeon's hands are not in direct contact with the tissue, but by means of long surgical instruments remotely controlled. As a result depth perception is reduced, hand eye coordination is distorted and haptic feedback is diminished. Fortunately, basic MIS skills (e.g. laparoscopic skills) can be acquired outside of the operating room (OR). This skills training results in a better performance during the actual surgical procedure. Regarding skills training outside of the OR, the value of various many laparoscopic exercises has been proven in terms of being construct valid, i.e. having the ability to discriminate surgeons of a different skill level. However, the scientific basic of other aspects of training facilities and the organisation of skills training is often lacking. Therefore, these subjects related to skills training outside of the OR are discussed in the first part of this thesis. The second part of this thesis investigates the value of the OSATS (Objective Structured Assessment of Technical Skills) for evaluation during actual surgery in the OR. Assessment with OSATS has been implemented while its value had only been studied in laboratory setting.

13

Outside of the operating room

In **chapter 2** the organisation of a basic surgical skills training is outlined for the Obstetrics and Gynaecology residency program in the Netherlands. In the light of this thesis, we specifically focussed on MIS (laparoscopy and hysteroscopy). Every resident is obliged to attend the same basic surgical skills course, intentionally during the first or second postgraduate year. One third of the course is spent on theory. For the complementary two thirds, the hands-on training, validated exercises are used with expert derived training goals based on time and precision. Furthermore, surgical skills are trained, expanded and assessed on simulators in the various teaching hospitals. Additional to this basic skills course, residents may attend advanced training courses focusing on laparoscopy and hysteroscopy. This organization guarantees a uniform introduction to MIS skills training for every resident. However, continued training and evaluation should be embedded in the curricula of the various teaching hospitals, and are the key to success of this approach after this uniform introduction.

Simulators are constantly being developed and improved. Virtual reality (VR) trainers and box trainers are available for skills training. Haptic feedback is naturally present in box trainers. Manufacturers of VR trainers have been looking for solutions to compensate the lack of haptic feedback. A possible solution is the addition of kinematic interaction between laparoscopic instruments and objects. **Chapter 3** presents a randomised controlled trial designed to determine which trainer model should be chosen for training. Additionally, it was determined whether the kinematic interaction in VR can replace haptic feedback of box trainers. Fifty novices were randomly assigned for training in a conventional VR setup, a VR environment

with additional kinematic interaction, a box trainer equivalent of both these setups or for a control group. An identical cylinder task was performed in all 4 training setups during 20 minutes. The effect was established by comparing the performance before and after training during a task that requires tissue handling. We found no improvement in the control group. The conventional VR group only improved in time, whereas VR with additional kinematic interaction and both box trainer groups improved in time, path length and motion in depth. We concluded that with respect to haptic feedback, box training models are superior to VR systems. However, additional kinematic interaction between instruments and objects can be a promising surrogate for haptic feedback in VR systems.

For box trainers, navigated and fixed camera systems are available. In **chapter 4** we compared the effect of a setup with a fixed camera with a setup with a camera navigated by the trainee and by an assistant on the performance during the initial learning phase. Sixty-nine right-handed medical students were randomized for one of the three camera setups. All had to perform eight trials of a task that requires hand eye coordination. We observed that time and the total path length of the three groups did not vary significantly along the eight trials. No significant difference was observed between the groups. A navigated camera offers theoretical advantages for the depth perception of the surgeon and allows practicing navigation skills, whereas a fixed setup allows solitary training. Therefore, combining training facilities with a fixed and a navigated setup would be superior.

In **chapter 5** we focused on the measurements used during assessment. We present a validation study of motion analysis parameters during an intracorporeal suturing task in a box trainer model. Novices, residents and laparoscopic experts performed three consecutive standardized intracorporeal sutures. Meanwhile, instrument movements were recorded using the TrEndo tracking device. The four investigated parameters (time, path length, motion in depth and motion smoothness) differed significantly in consecutive level of experience. Therefore, the construct validity has been proved for these parameters for the laparoscopic suturing task using a box trainer. Besides, an expert level has been set for training and assessment purposes. Furthermore, the addition of economy of movement to time to complete the task has the potential to refine acquisition of skills.

The purpose of the study presented in **chapter 6** was to test the retention of basic laparoscopic skills on a box trainer one year after a short training program. Eight medical students without prior experience underwent baseline testing, followed by five weekly training sessions and a final test. During each of seven sessions, they performed five tasks on an inanimate box trainer. Scores were calculated by adding up the time to completion of the task with penalty points, consequently rewarding speed and precision. The sum score was the total of the scores of the five tasks. One year later, seven of them underwent retention testing. The final test results were compared with retention test results as a measure of durability of acquired skills. Novices' scores did not worsen significantly for four out of five tasks (placing a pipe cleaner, placing beads, cutting a circle and knot tying). However, deterioration was observed in the performance of stretching a rubber band, as well as in the sum score. In conclusion, basic laparoscopic skills acquired during a short training program merely sustain over time. However, on-going practice is advisable, especially to preserve tissue-handling skills, since these may be the first to deteriorate.

In **Chapter 7** we developed an international and consensus based set of quality criteria for a skills laboratory for training MIS, including aspects of the design of the laboratory and the training curriculum. Three quality domains for skills laboratory were defined; Personnel and Resources, Trainee motivation and training Curriculum. A list of consensus-based criteria, 9 items per domain, was made. Additionally, 23 worldwide experts in MIS were asked to rate each item on a 0 to 3 scale in level of importance. In the domain Personnel and resources, the presence of a box trainer, a laparoscopic expert and the availability of financial resources were considered the most important. In the domain Trainee motivation, mandatory training supervised by laparoscopic experts were considered the most important. In the domain Curriculum, the presence of a structured skills curriculum, dedicated time for skills training, maintenance of the skills and a yearly evaluation of the progress were considered the most important factors. This rating list can be used when setting up a skills laboratory, but also for verifying the quality of an existing laboratory. From there, the focus for new developments can be chosen.

In the operating room

Chapter 8 is a validation study of the OSATS for intraoperative use. Nine residents participated. We mapped individual learning curves of residents using OSATS scores as a measure of performance. We tested the hypothesis that with increasing experience within a certain procedure the OSATS score will raise. This hypothesis was confirmed. The OSATS score significantly increases by an average of 1.10 points per assessed procedure. We noticed that the median OSATS scores among the 21 supervisors ranged from 18 to 30. We concluded that intraoperative assessment with OSATS have construct validity. Furthermore, the individual learning curves enable insight into individual progression.

In **Chapter 9** we evaluated additionally relevant issues regarding the implementation of the OSATS as an intraoperative assessment tool by collecting assessment forms and performing a survey among users of the OSATS. We found that 28 of 30 points is the median OSATS score at which a resident can perform a procedure autonomously. The intraclass correlation coefficient between resident and supervisor is 0.78 which indicates substantial agreement. Moreover, residents and supervisors do not judge the OSATS to be very objective. A first step has been made towards setting benchmark criteria for using OSATS for authorization. However, the limited objectivity of this instrument should be taken into account. Furthermore, it is advisable to focus attention on other competencies of a surgeon too.

It is frequently suggested that MIS is harder to acquire than conventional surgery. To test this hypothesis, residents' learning curves of both surgical skills were compared in **chapter 10**. Nine residents collected a total of 319 OSATS during their three-month rotation in gynaecological surgery. Learning curves for MIS (laparoscopic and hysteroscopic) and conventional surgery (open abdominal and vaginal) were compared. The MIS curve revealed to be steeper than the conventional curve (1.77 versus 0.75 OSATS points per assessed procedure). We concluded that basic MIS procedures do not seem harder to acquire during residency than conventional surgical procedures. This may have resulted from the incorporation of structured MIS training programs in residency. Hopefully, this will lead to a more successful implementation of the advanced MIS procedures.

In **chapter 11** the combined results of the aforementioned chapters are discussed in a broader perspective. In conclusion, a structured curriculum in which (minimally invasive) surgical skills are trained and evaluated is indispensible during surgical specialty training. However, a careful consideration should be made about the contains of this curriculum, with reaching surgical proficiency as the ultimate goal.

13

SAMENVATTING

In **hoofdstuk 1** wordt de achtergrond beschreven van de studies die gepresenteerd worden in dit proefschrift. Minimaal invasieve chirurgie (MIC) is ontwikkeld als alternatieve benaderingswijze voor ingrepen die met een grote incisie (zoals bij een laparotomie) gepaard gaan en heeft als doel zo min mogelijk weefselschade te veroorzaken. Het gebruik van deze techniek leidt daarom tot een sneller herstel van een patiënt en geeft cosmetisch een fraaier resultaat. De vaardigheden die vereist zijn om deze vorm van chirurgie veilig en efficiënt uit te voeren vormen echter een grote uitdaging. Het operatiegebied wordt immers indirect in beeld gebracht via een cameraprojectie op een tweedimensionaal scherm. Daarnaast heeft de (gehandschoende) hand van de chirurg geen direct contact met de weefsels, maar op afstand via instrumenten die buiten het lichaam van een patiënt worden bediend. Het resultaat is dat er minder dieptewaarneming is, dat oog-handcoördinatie bemoeilijkt wordt en dat er minder gevoelsterugkoppeling vanuit de weefsels is. Gelukkig kunnen minimaal invasieve (bijvoorbeeld laparoscopische) chirurgische vaardigheden buiten de operatiekamer aangeleerd worden. Deze training resulteert in een verbeterde uitvoering van de echte chirurgische procedures in de praktijk. De waarde van een veelheid aan oefeningen buiten de operatiekamers is bewezen doordat aangetoond is dat chirurgen die in ervaringsniveau van elkaar verschillen, ook verschillend presteren op de oefeningen. Echter, van andere aspecten van trainingsopstellingen en van de organisatie van de chirurgische vaardigheidstraining in een opleidingscurriculum is de wetenschappelijke basis vaak afwezig. Het eerste gedeelte van dit proefschrift is daarom gericht op onbeantwoorde vragen aangaande vaardigheidstraining buiten de operatiekamer. Het tweede gedeelte onderzoekt de waarde van de OSATS (objectieve gestructureerde beoordeling van technische vaardigheden) om chirurgische vaardigheden in de operatiekamer te evalueren. Deze vorm van beoordelen was al geïmplementeerd, terwijl de waarde van OSATS alleen voor proefopstellingen buiten de operatiekamer bewezen was.

Buiten de operatiekamer

In **hoofdstuk 2** wordt de organisatie van de basale chirurgische vaardigheidstraining tijdens de opleiding Obstetrie en Gynaecologie in Nederland uiteengezet. Omdat MIC het onderwerp van dit proefschrift is, is dit hoofdstuk gericht op laparoscopie en hysteroscopie. Elke arts in opleiding tot specialist (AIOS) is verplicht dezelfde cursus *basale chirurgische vaardigheden* te volgen, bij voorkeur gedurende de eerste twee jaren van de opleiding. Een derde van de tijd van de cursus wordt besteed aan theoretische kennis. De rest wordt besteed aan praktische vaardigheidstraining, waarbij gevalideerde oefeningen worden gebruikt. Het einddoel is daarbij het bereiken van het niveau van een expert qua snelheid en precisie. Vervolgens moeten deze vaardigheden worden onderhouden, worden uitgebreid en met regelmaat worden getoetst op simulatoren in de verschillende opleidingscentra verspreid over het land. Als aanvulling op deze basiscursus kunnen AIOS cursussen voor gevorderden volgen in hysteroscopie en/of laparoscopie. Deze organisatie garandeert een uniforme introductie in vaardigheidstraining in MIC voor elke AIOS. Echter, vervolgtraining en evaluatie moeten ingebed worden in de curricula van de verschillende opleidingsziekenhuizen. Dit is essentieel voor het succes van deze aanpak.

Simulatoren worden continu ontwikkeld en verbeterd. Er bestaan oefenboxen en *virtual reality* (VR) trainers. In oefenboxen is terugkoppeling van krachten van de weefsels van nature aanwezig. Fabrikanten van VR trainers hebben naar oplossingen gezocht om de afwezigheid van krachtsterugkoppeling van de weefsels (of objecten) in hun trainers te compenseren. Een voorbeeld van een dergelijke oplossing is het toevoegen van een extra bewegingsinteractie tussen instrumenten en voorwerpen. **Hoofdstuk 3** beschrijft een gerandomiseerd gecontroleerd onderzoek om te ontdekken welk trainingsmodel het beste gebruikt kan worden. Daarnaast wordt bepaald of de beschreven bewegingsinteractie in VR de krachtsterugkoppeling van oefenboxen kan vervangen. Vijftig studenten werden middels loting verdeeld over vijf groepen: een conventionele VR opstelling, een VR opstelling met de bewegingsinteractie, de twee oefenbox-equivalenten van deze opstellingen en een controle groep. In elk van de vier trainingsopstellingen werd een identieke taak uitgevoerd om cilinders op elkaar te stapelen gedurende 20 minuten. Het effect van deze training werd vastgesteld door een pre- en post-test waarbij weefselgevoel een vereiste is. De controlegroep presteerde niet beter tijdens de post-test. De groep die op de conventionele VR opstelling getraind had verbeterde alleen in tijd, terwijl de groep van de VR met de bewegingsinteractie ook verbeterde in de bewegingsparameters pad-lengte en beweging in diepterichting. Wij concludeerden dat oefenboxen superieur zijn boven VR trainers wat betreft haptische terugkoppeling. Echter, de toevoeging van bewegingsinteractie tussen instrumenten en objecten is een veelbelovend surrogaat voor haptische terugkoppeling in VR systemen.

Er zijn zowel gefixeerde als navigeerbare camerasystemen beschikbaar voor oefenboxen. In **hoofdstuk 4** onderzochten wij het effect op het eerste gedeelte van de leercurve van een opstelling met een gefixeerde camera versus een opstelling met een camera die genavigeerd wordt door een assistent versus een opstelling waarbij de camera wordt genavigeerd door degene die zelf de oefening uitvoert. Negenenzestig rechtshandige studenten werden middels loting bij één van de drie cameraopstellingen ingedeeld. Allen moesten zij acht pogingen doen om een taak uit te voeren die oog-handcoördinatie vereist. Er werden geen significant verschillen gemeten in tijd of in bewegingsparameters tijdens de acht pogingen. Een navigeerbare camera biedt theoretische voordelen voor het diepte-inzicht en bovendien worden de navigatie-vaardigheden getraind. Aan de andere kant maakt een opstelling met een gefixeerde camera training mogelijk zonder dat er anderen bij aanwezig zijn. Omdat wij geen verschil aantoonden in de leercurves, lijkt de combinatie van een gefixeerd en een navigeerbaar camerasysteem de beste optie.

In **hoofdstuk 5** richtten wij ons op de uitkomstmaten tijdens beoordeling. Wij presenteren een validatie studie van bewegingsparameters tijdens het leggen van een laparoscopische hechting in een oefenbox. Studenten, AIOS en laparoscopische experts legden drie laparoscopische hechtingen in een oefenbox. Terwijl zij dit deden werden de bewegingen die de tip van hun instrumenten maakten geregistreerd met de TrEndo. De vier onderzochte parameters (tijd, pad-lengte, beweging in diepterichting en een maat voor hoe vloeiend de bewegingen verlopen) verschilden significant voor alle drie de ervaringsniveaus. Daarmee werd de construct-validiteit van deze parameters bewezen voor de laparoscopische hechttaak in een oefenbox. Bovendien werd het niveau van een expert bepaald. Dit niveau kan nu gebruikt worden voor trainings- en beoordelingsdoeleinden. Daarnaast heeft het toevoegen van parameters van de efficiëntie van beweging de potentie om het aanleren van vaardigheden te verfijnen.

Het doel van de studie die gepresenteerd wordt in **hoofdstuk 6** was om te onderzoeken in hoeverre basale laparoscopische vaardigheden die tijdens een korte training op een oefenbox zijn geleerd beklijven na verloop van een jaar. Acht medisch studenten zonder laparoscopische ervaring ondergingen een pre-test, gevolgd door vijf trainingssessies en een eind-test. Tijdens deze in totaal zeven sessies voerden zij vijf taken uit op een oefenbox. Hun prestatie werd gescoord door de tijd te vermeerderen met strafpunten. Zo werden zowel snelheid als precisie beloond. De totale score was de som van de scores van de vijf taken. Een jaar later ondergingen zeven van hen een test om het beklijven van de verkregen vaardigheden te meten zonder tussenliggende training. Hun score verslechterde niet significant voor vier taken (pijpenrager manoeuvreren, kralen plaatsen, cirkel knippen en laparoscopisch hechten). Echter, een verslechtering werd geobserveerd voor de taak om een elastiek op spanning te plaatsen rond een aantal spijkers. Ook de totale score verslechterde. Wij concludeerden daarom dat vaardigheden verkregen tijdens een kort trainingsprogramma grotendeels beklijven. Echter, onderhoud is verstandig, vooral met het oog op het behouden van weefselgevoel, want dat lijkt de eerste vaardigheid te zijn die men verleerd.

In **hoofdstuk 7** hebben wij een internationale set van kwaliteitscriteria ontwikkeld voor trainingscentra voor MIC. Deze criteria gaan zowel over het opzetten van een centrum als de organisatie van vaardigheidstraining in een opleidingscurriculum. Drie kwaliteitsdomeinen werden gedefinieerd: Personeel en materiaal, Motivatie en Curriculum. Vervolgens werd een lijst opgesteld met negen criteria per domein. Vervolgens werden 23 internationale experts op het gebied van MIC gevraagd het belang van elk van de criteria te scoren op een schaal van 0 tot 3. Binnen het domein Personeel en materiaal werden de aanwezigheid van een oefenbox, van een laparoscopische expert en de beschikbaarheid van geld het meest belangrijk gevonden. Binnen het domein Motivatie werden verplichte trainingen, gesuperviseerd door een expert het meest van belang geacht. Binnen het domein Curriculum werden de aanwezigheid van een gestructureerd curriculum met ingeroosterde trainingstijd, aandacht voor onderhoud van verkregen vaardigheden en jaarlijkse evaluatie hiervan het meest belangrijk gevonden. Deze gewogen lijst van kwaliteitscriteria kan zowel gebruikt worden bij het opzetten van een trainingscentrum, als bij het beoordelen van de kwaliteit van bestaande trainingsfaciliteiten. Een dergelijke beoordeling kan helpen bij de keuze van nieuwe ontwikkelingen.

In de operatiekamer

Hoofdstuk 8 is een validatie-studie van de OSATS voor gebruik in de operatiekamer. Negen AIOS namen deel. We gaven per AIOS een individuele leercurve weer door de totale OSATS score als maat voor prestatie te nemen. Wij testten de hypothese dat wanneer de ervaring binnen een bepaalde chirurgische ingreep toeneemt ook de OSATS score stijgt. Deze hypothese werd bevestigd. De OSATS score stijgt met gemiddeld 1.10 punten per volgende beoordeelde ingreep. Een kanttekening is dat de mediane score die elk van de 21 supervisoren gaf uiteenliep van 18 tot 30. Deze studie vormt een aanwijzing dat de OSATS construct-valide zijn.

In **hoofdstuk 9** onderzochten wij aspecten van OSATS die naast construct-validiteit van belang zijn voor de implementatie van OSATS in de operatiekamer. Hiertoe verzamelden wij OSATS beoordelingen en namen een vragenlijst af. Het bleek dat bij een score van 28 (30 is het maximum) een AIOS door de supervisor in staat wordt geacht een procedure zelfstandig uit te voeren. Er was sprake van substantiële overeenstemming tussen supervisor en AIOS. Daarentegen gaven zowel AIOS als supervisoren aan de objectiviteit van de OSATS te betwisten. Met deze studie is een eerste stap gezet om de OSATS te ijken als instrument voor autorisatie. Echter, de beperkte objectiviteit van de OSATS moet in acht genomen worden. Daarnaast moet er ook aandacht besteed worden aan andere competenties van een chirurg.

Het wordt vaak gesuggereerd dat MIC moeilijker aan te leren is dan conventionele chirurgie. Om deze hypothese te testen werden de leercurven van beide vormen van chirurgie met elkaar vergeleken in **hoofdstuk 10**. Negen AIOS verzamelden in totaal 319 OSATS tijdens hun drie maanden durende stage gynaecologische chirurgie. Leercurves voor MIC (laparoscopie en hysteroscopie) en conventionele chirurgie (laparotomie en vaginale chirurgie) werden met elkaar vergeleken. Het bleek dat de leercurve voor MIC steiler was dan de curve voor conventionele chirurgie (1.77 versus 0.75 OSATS punten per beoordeelde ingreep). We concludeerden dat de basale MIC procedures niet moeilijker aan te leren lijken dan conventionele procedures tijdens de opleiding tot gynaecoloog. Dit zou het resultaat kunnen zijn van de gestructureerde invoering van MIC vaardigheidstraining tijdens de opleiding. Hopelijk zal dit ook leiden tot een betere implementatie van de geavanceerde MIC procedures.

In **hoofdstuk 11** worden de resultaten van de voorgaande hoofdstukken besproken en bediscussieerd. Kort samengevat is een gestructureerd curriculum waarbinnen (minimaal invasieve) chirurgische vaardigheden worden getraind en geëvalueerd tegenwoordig niet meer weg te denken uit de opleiding tot snijdend specialist. Welke elementen aan een dergelijk curriculum moeten worden toegevoegd en welke moeten worden afgeschaft blijft een afweging van essentieel belang. Het ultieme doel is een vakbekwame chirurg.

CHAPTER 14 | ADDENDUM

LITERATURE
AUTHOR AFFILIATIONS
ABOUT THE AUTHOR
PUBLICATIONS
DANKWOORD

LITERATURE

Aggarwal, R., Grantcharov, T., Moorthy, K., Milland, T., & Darzi, A. (2008). Toward feasible, valid, and reliable video-based assessments of technical surgical skills in the operating room. *Ann.Surg., 247,* 372-379.

Aggarwal, R., Hance, J., Undre, S., Ratnasothy, J., Moorthy, K., Chang, A. et al. (2006). Training junior operative residents in laparoscopic suturing skills is feasible and efficacious. *Surgery, 139,* 729-734.

Aggarwal, R., Moorthy, K., & Darzi, A. (2004). Laparoscopic skills training and assessment. *Br.J.Surg., 91,* 1549-1558.

Altgassen, C., Michels, W., & Schneider, A. (2004). Learning laparoscopic-assisted hysterectomy. *Obstet.Gynecol., 104,* 308-313.

Ames, C., Frisella, A. J., Yan, Y., Shulam, P., & Landman, J. (2006). Evaluation of laparoscopic performance with alteration in angle of vision. *J.Endourol., 20,* 281-283.

Anastakis, D. J., Regehr, G., Reznick, R. K., Cusimano, M., Murnaghan, J., Brown, M. et al. (1999). Assessment of technical skills transfer from the bench training model to the human model. *Am.J.Surg., 177,* 167-170.

Arthur W Jr, B. W. J. S. P. M. T. (1998). (1998) Factors that Influence Skill Decay and Retention: A Quantitative Review and Analysis. Human Performance 11:57-101. Human Performance 11, 57-101.

Basdogan, C., De, S., Kim, J., Muniyandi, M., Kim, H., & Srinivasan, M. A. (2004). Haptics in minimally invasive surgical simulation and training. *IEEE Comput.Graph.Appl., 24,* 56-64.

Basdogan, C., Sedef, M., Harders, M., & Wesarg, S. (2007). VR-based simulators for training in minimally invasive surgery. *IEEE Comput.Graph. Appl., 27,* 54-66.

Bashankaev, B., Baido, S., & Wexner, S. D. (2011). Review of available methods of simulation training to facilitate surgical education. *Surg. Endosc., 25,* 28-35.

Beard, J. D. (2007). Assessment of surgical competence. *Br.J.Surg., 94,* 1315-1316.

Bennett, A., Birch, D. W., Menzes, C., Vizhul, A., & Karmali, S. (2011). Assessment of medical student laparoscopic camera skills and the impact of formal camera training. *Am.J.Surg., 201,* 650-654.

Bholat, O. S., Haluck, R. S., Murray, W. B., Gorman, P. J., & Krummel, T. M. (1999). Tactile feedback is present during minimally invasive surgery. *J.Am.Coll.Surg., 189,* 349-355.

Bijen, C. B., Briet, J. M., de Bock, G. H., Arts, H. J., Bergsma-Kadijk, J. A., & Mourits, M. J. (2009). Total laparoscopic hysterectomy versus abdominal hysterectomy in the treatment of patients with early stage endometrial cancer: a randomized multi center study. *BMC.Cancer, 9,* 23.

Blanchard, M. H., Amini, S. B., & Frank, T. M. (2004). Impact of work hour restrictions on resident case experience in an obstetrics and gynecology residency program. *Am.J.Obstet. Gynecol., 191,* 1746-1751.

Bodle, J. F., Kaufmann, S. J., Bisson, D., Nathanson, B., & Binney, D. M. (2008). Value and face validity of objective structured assessment of technical skills (OSATS) for work based assessment of surgical skills in obstetrics and gynaecology. *Med.Teach., 30,* 212-216.

Botden, S. M., Torab, F., Buzink, S. N., & Jakimowicz, J. J. (2008). The importance of haptic feedback in laparoscopic suturing training and the additive value of virtual reality simulation. *Surg.Endosc., 22,* 1214-1222.

Brolmann, H. A., Vervest, H. A., & Heineman, M. J. (2001). Declining trend in major gynaecological surgery in The Netherlands during 1991-1998. Is there an impact on surgical skills and innovative ability? *BJOG., 108,* 743-748.

Cadeddu, J. A., Wolfe, J. S., Jr., Nakada, S., Chen, R., Shalhav, A., Bishoff, J. T. et al. (2001). Complications of laparoscopic procedures after concentrated training in urological laparoscopy. *J.Urol., 166,* 2109-2111.

Chang, L., Petros, J., Hess, D. T., Rotondi, C., & Babineau, T. J. (2007). Integrating simulation into a surgical residency program: is voluntary participation effective? *Surg.Endosc., 21,* 418-421.

Chen, C. C., Korn, A., Klingele, C., Barber, M. D., Paraiso, M. F., Walters, M. D. et al. (2010). Objective assessment of vaginal surgical skills. *Am.J.Obstet.Gynecol.*.

Chmarra, M. K., Bakker, N. H., Grimbergen, C. A., & Dankelman, J. (2006). TrEndo, a device for tracking minimally invasive surgical instruments in training setups. *Sensors and Actuators A-Physical, 126,* 328-334.

Chmarra, M. K., Dankelman, J., van den Dobbelsteen, J. J., & Jansen, F. W. (2008). Force feedback and basic laparoscopic skills. *Surg. Endosc., 22,* 2140-2148.

Chmarra, M. K., Grimbergen, C. A., & Dankelman, J. (2007). Systems for tracking minimally invasive surgical instruments. *Minim.Invasive.Ther.Allied Technol., 16,* 328-340.

&

Clasen, K., Camus, M., Tournaye, H., & Devroey, P. (1997). Ectopic pregnancy: let's cut! Strict laparoscopic approach to 194 consecutive cases and review of literature on alternatives. *Hum. Reprod., 12,* 596-601.

Conrad, J., Shah, A. H., Divino, C. M., Schluender, S., Gurland, B., Shlasko, E. et al. (2006). The role of mental rotation and memory scanning on the performance of laparoscopic skills: a study on the effect of camera rotational angle. *Surg. Endosc., 20,* 504-510.

Cotin, S., Stylopoulos, N., Ottensmeyer, M., Neumann, P., Rattner, D., & Dawson, S. (2002). Metrics for laparoscopic Skills Trainers: The weakest link! In *Lecture Notes in Computer Science* (pp. 35-43).

Cuschieri, A., Francis, N., Crosby, J., & Hanna, G. B. (2001). What do master surgeons think of surgical competence and revalidation? *Am.J.Surg., 182,* 110-116.

Dagash, H., Chowdhury, M., & Pierro, A. (2003). When can I be proficient in laparoscopic surgery? A systematic review of the evidence. *J.Pediatr. Surg., 38,* 720-724.

Darzi, A. & Mackay, S. (2002). Recent advances in minimal access surgery. *BMJ, 324,* 31-34.

Darzi, A., Smith, S., & Taffinder, N. (1999). Assessing operative skill. Needs to become more objective. *BMJ, 318,* 887-888.

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA, 296,* 1094-1102.

Derossis, A. M., Bothwell, J., Sigman, H. H., & Fried, G. M. (1998). The effect of practice on performance in a laparoscopic simulator. *Surg. Endosc., 12,* 1117-1120.

Dunkin, B., Adrales, G. L., Apelgren, K., & Mellinger, J. D. (2007). Surgical simulation: a current review. *Surg.Endosc., 21,* 357-366.

Elwyn, G., O'Connor, A., Stacey, D., Volk, R., Edwards, A., Coulter, A. et al. (2006). Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ, 333,* 417.

Emam, T. A., Hanna, G., & Cuschieri, A. (2002). Comparison of orthodox versus off-optical axis endoscopic manipulations. *Surg.Endosc., 16,* 401-405.

ESGE 2009: http://www.esge.org

Eva, K. W. & Regehr, G. (2005). Self-assessment in the health professions: a reformulation and research agenda. *Acad.Med., 80,* S46-S54.

Feldman, L. S., Hagarty, S. E., Ghitulescu, G., Stanbridge, D., & Fried, G. M. (2004a). Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. *J.Am.Coll.Surg., 198,* 105-110.

Feldman, L. S., Sherman, V., & Fried, G. M. (2004b). Using simulators to assess laparoscopic competence: ready for widespread use? *Surgery, 135,* 28-42.

Forde, K. A. (1993). Endosurgical training methods: is it surgical training that is out of control? *Surg.Endosc., 7,* 71-72.

Frank, J. R. & Danoff, D. (2007). The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med. Teach., 29,* 642-647.

Fried, G. M., Derossis, A. M., Bothwell, J., & Sigman, H. H. (1999). Comparison of laparoscopic performance in vivo with performance measured in a laparoscopic simulator. *Surg. Endosc., 13,* 1077-1081.

Fried, G. M., Feldman, L. S., Vassiliou, M. C., Fraser, S. A., Stanbridge, D., Ghitulescu, G. et al. (2004). Proving the value of simulation in laparoscopic surgery. *Ann.Surg., 240,* 518-525.

Gallagher, A. G., McClure, N., McGuigan, J., Ritchie, K., & Sheehy, N. P. (1998). An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy, 30,* 617-620.

Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G. et al. (2005). Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann.Surg., 241,* 364-372.

Ginsburg, S., Regehr, G., Hatala, R., McNaughton, N., Frohna, A., Hodges, B. et al. (2000). Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad.Med., 75,* S6-S11.

Goff, B., Mandel, L., Lentz, G., Vanblaricom, A., Oelschlager, A. M., Lee, D. et al. (2005). Assessment of resident surgical skills: is testing feasible? *Am.J.Obstet.Gynecol., 192,* 1331-1338.

Goff, B. A., Lentz, G. M., Lee, D., Houmard, B., & Mandel, L. S. (2000). Development of an objective structured assessment of technical skills for obstetric and gynecology residents. *Obstet.Gynecol., 96,* 146-150.

Goff, B. A., Nielsen, P. E., Lentz, G. M., Chow, G. E., Chalmers, R. W., Fenner, D. et al. (2002). Surgical skills assessment: a blinded examination of obstetrics and gynecology residents. *Am.J.Obstet.Gynecol., 186,* 613-617.

Gordon, M. J. (1992). Self-assessment programs and their implications for health professions training. *Acad.Med., 67,* 672-679.

Gould, J. C. (2006). Building a laparoscopic surgical skills training laboratory: resources and support. *JSLS., 10,* 293-296.

Grantcharov, T. P., Kristiansen, V. B., Bendix, J., Bardram, L., Rosenberg, J., & Funch-Jensen, P. (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br.J.Surg., 91,* 146-150.

Grober, E. D., Hamstra, S. J., Wanzel, K. R., Reznick, R. K., Matsumoto, E. D., Sidhu, R. S. et al. (2004a). Laboratory based training in urological microsurgery with bench model simulators: a randomized controlled trial evaluating the durability of technical skill. *J.Urol., 172,* 378-381.

Grober, E. D., Hamstra, S. J., Wanzel, K. R., Reznick, R. K., Matsumoto, E. D., Sidhu, R. S. et al. (2004b). The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. *Ann.Surg., 240,* 374-381.

Haluck, R. S. & Krummel, T. M. (2000). Computers and virtual reality for surgical education in the 21st century. *Arch.Surg., 135,* 786-792.

Hamilton, E. C., Scott, D. J., Fleming, J. B., Rege, R. V., Laycock, R., Bergen, P. C. et al. (2002). Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills. *Surg.Endosc., 16,* 406-411.

Hammond, I. & Karthigasu, K. (2006). Training, assessment and competency in gynaecologic surgery. *Best.Pract.Res.Clin.Obstet.Gynaecol., 20,* 173-187.

Hammoud, M. M., Nuthalapaty, F. S., Goepfert, A. R., Casey, P. M., Emmons, S., Espey, E. L. et al. (2008). To the point: medical education review of the role of simulators in surgical training. *Am.J.Obstet.Gynecol., 199,* 338-343.

Hanna, G. B., Shimi, S. M., & Cuschieri, A. (1998). Task performance in endoscopic surgery is influenced by location of the image display. *Ann. Surg., 227,* 481-484.

Haveran, L. A., Novitsky, Y. W., Czerniach, D. R., Kaban, G. K., Taylor, M., Gallagher-Dorval, K. et al. (2007). Optimizing laparoscopic task efficiency: the role of camera and monitor positions. *Surg.Endosc., 21,* 980-984.

Heemskerk, J., Zandbergen, R., Maessen, J. G., Greve, J. W., & Bouvy, N. D. (2006). Advantages of advanced laparoscopic systems. *Surg. Endosc., 20,* 730-733.

Hiemstra, E., Kolkman, W., & Jansen, F. W. (2008). Skills training in minimally invasive surgery in Dutch obstetrics and gynecology residency curriculum. *Gynecol.Surg., 5,* 321-325.

Hiemstra E, Jansen F.W. (2010). OSATS Guideline. *Dutch.J for.Med.Educ.*, 29; 5: 135-141

Hiemstra, E., Kolkman, W., Wolterbeek, R., Trimbos, B., & Jansen, F. W. (2011). Value of an objective assessment tool in the operating room. *Can.J.Surg., 54,* 116-122.

Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. *Acad.Med., 76,* S87-S89.

Horeman, T., Rodrigues, S. P., Jansen, F. W., Dankelman, J., & van den Dobbelsteen, J. J. (2010). Force measurement platform for training and assessment of laparoscopic skills. *Surg. Endosc., 24,* 3102-3108.

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

IGZ 2007: http://www.igz.nl

Jansen F.W. & Trimbos-Kemper G.C.M (2008). *Hysteroscopic surgery, the basis*.

Jansen, F. W. & Trimbos-Kemper G.C.M (2006). *Laparoscopy, the basis*.

Jordan, J. A., Gallagher, A. G., McGuigan, J., McGlade, K., & McClure, N. (2000). A comparison between randomly alternating imaging, normal laparoscopic imaging, and virtual reality training in laparoscopic psychomotor skill acquisition. *Am.J.Surg., 180,* 208-211.

Kingston, A., Abbott, J., Lenart, M., & Vancaillie, T. (2004). Hysteroscopic training: the butternut pumpkin model. *J.Am.Assoc.Gynecol.Laparosc., 11,* 256-261.

Kolkman, W., Engels, L. E., Smeets, M. J., & Jansen, F. W. (2007a). Teach the teachers: an observational study on mentor traineeship in gynecological laparoscopic surgery. *Gynecol. Obstet.Invest, 64,* 1-7.

Kolkman, W., van de Put, M. A., Van den Hout, W. B., Trimbos, J. B., & Jansen, F. W. (2007b). Implementation of the laparoscopic simulator in a gynecological residency curriculum. *Surg. Endosc., 21,* 1363-1368.

Kolkman, W., Wolterbeek, R., & Jansen, F. W. (2005). Gynecological laparoscopy in residency training program: Dutch perspectives. *Surg. Endosc., 19,* 1498-1502.

Kolkman, W., Wolterbeek, R., & Jansen, F. W. (2006). Implementation of advanced laparoscopy into daily gynecologic practice: difficulties and solutions. *J.Minim.Invasive.Gynecol., 13,* 4-9.

Kolkman, W., van de Put, M. A. J., & Jansen, F. W. (2008). Laparoscopic simulator: construct validity and establishing performance standards for residency training. Gyn Surg; 5:109-114. *Gynecological Surgery.*

Korndorffer, J. R., Jr., Dunne, J. B., Sierra, R., Stefanidis, D., Touchard, C. L., & Scott, D. J. (2005a). Simulator training for laparoscopic suturing using performance goals translates to the operating room. *J.Am.Coll.Surg., 201,* 23-29.

&

Korndorffer, J. R., Jr., Hayes, D. J., Dunne, J. B., Sierra, R., Touchard, C. L., Markert, R. J. et al. (2005b). Development and transferability of a cost-effective laparoscopic camera navigation simulator. *Surg.Endosc., 19,* 161-167.

Korndorffer, J. R., Jr., Scott, D. J., Sierra, R., Brunner, W. C., Dunne, J. B., Slakey, D. P. et al. (2005c). Developing and testing competency levels for laparoscopic skills training. *Arch.Surg., 140,* 80-84.

Korndorffer, J. R., Jr., Stefanidis, D., & Scott, D. J. (2006). Laparoscopic skills laboratories: current assessment and a call for resident training standards. *Am.J.Surg., 191,* 17-22.

Kothari, S. N., Kaplan, B. J., DeMaria, E. J., Broderick, T. J., & Merrell, R. C. (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, 12,* 167-173.

Kruger, J., Gordon, C. L., & Kuban, J. (2006). Intentions in teasing: when "just kidding" just isn't good enough. *J.Pers.Soc.Psychol., 90,* 412-425.

Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. In *Biometrics* (pp. 159-174).

Larsen, C. R., Grantcharov, T., Aggarwal, R., Tully, A., Sorensen, J. L., Dalsgaard, T. et al. (2006). Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surg.Endosc., 20,* 1460-1466.

Larsen, C. R., Soerensen, J. L., Grantcharov, T. P., Dalsgaard, T., Schouenborg, L., Ottosen, C. et al. (2009). Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ, 338,* b1802.

Lentz, G. M., Mandel, L. S., Lee, D., Gardella, C., Melville, J., & Goff, B. A. (2001). Testing surgical skills of obstetric and gynecologic residents in a bench laboratory setting: validity and reliability. *Am.J.Obstet.Gynecol., 184,* 1462-1468.

Loh, F. H., Hameed, N., & Ng, S. C. (2002). The impact of minimal access surgery on gynaecological surgery in a university gynaecological unit over a 10-year period from 1991 to 2000. *Singapore Med.J., 43,* 177-181.

Lynn, D. J., Holzer, C., & O'Neill, P. (2006). Relationships between self-assessment skills, test performance, and demographic variables in psychiatry residents. *Adv.Health Sci.Educ.Theory. Pract., 11,* 51-60.

Maagaard, M., Sorensen, J. L., Oestergaard, J., Dalsgaard, T., Grantcharov, T. P., Ottesen, B. S. et al. (2011). Retention of laparoscopic procedural skills acquired on a virtual-reality surgical trainer. *Surg.Endosc., 25,* 722-727.

MacRae, H. M., Satterthwaite, L., & Reznick, R. K. (2008). Setting up a surgical skills center. *World J.Surg., 32,* 189-195.

Madan, A. K. & Frantzides, C. T. (2007). Prospective randomized controlled trial of laparoscopic trainers for basic laparoscopic skills acquisition. *Surg.Endosc., 21,* 209-213.

Mahmood, T. & Darzi, A. (2004). The learning curve for a colonoscopy simulator in the absence of any feedback: no feedback, no learning. *Surg. Endosc., 18,* 1224-1230.

Mandel, L. P., Lentz, G. M., & Goff, B. A. (2000). Teaching and evaluating surgical skills. *Obstet. Gynecol., 95,* 783-785.

Mandel, L. S., Goff, B. A., & Lentz, G. M. (2005). Self-assessment of resident surgical skills: is it feasible? *Am.J.Obstet.Gynecol., 193,* 1817-1822.

Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C. et al. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *Br.J.Surg., 84,* 273-278.

Martin, M., Vashisht, B., Frezza, E., Ferone, T., Lopez, B., Pahuja, M. et al. (1998). Competency-based instruction in critical invasive skills improves both resident performance and patient safety. *Surgery, 124,* 313-317.

Matern, U., Faist, M., Kehl, K., Giebmeyer, C., & Buess, G. (2005). Monitor position in laparoscopic surgery. *Surg.Endosc., 19,* 436-440.

McClusky, D. A., III & Smith, C. D. (2008). Design and development of a surgical skills simulation curriculum. *World J.Surg., 32,* 171-181.

Medeiros, L. R., Stein, A. T., Fachel, J., Garry, R., & Furness, S. (2008). Laparoscopy versus laparotomy for benign ovarian tumor: a systematic review and meta-analysis. *Int.J.Gynecol. Cancer, 18,* 387-399.

Moore, M. J. & Bennett, C. L. (1995). The learning curve for laparoscopic cholecystectomy. The Southern Surgeons Club. *Am.J.Surg., 170,* 55-59.

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

Moschos, E. & Coleman, R. L. (2004). Acquiring laparoscopic skill proficiency: does orientation matter? *Am.J.Obstet.Gynecol., 191,* 1782-1787.

Moulton, C. A., Dubrowski, A., MacRae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Ann.Surg., 244,* 400-409.

Munz, Y., Almoudaris, A. M., Moorthy, K., Dosis, A., Liddle, A. D., & Darzi, A. W. (2007). Curriculum-based solo virtual reality training for laparoscopic intracorporeal knot tying: objective

&

assessment of the transfer of skill from virtual reality to reality. *Am.J.Surg., 193,* 774-783.

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

Najmaldin, A. (2007). Karl Storz Lecture. Skills training in pediatric minimal access surgery. *J.Pediatr.Surg., 42,* 284-289.

Navez, B. & Penninckx, F. (1999). Laparoscopic training: results of a Belgian survey in trainees. Belgian Group for Endoscopic Surgery (BGES). *Acta Chir Belg., 99,* 53-58.

Newmark, J., Dandolu, V., Milner, R., Grewal, H., Harbison, S., & Hernandez, E. (2007). Correlating virtual reality and box trainer tasks in the assessment of laparoscopic surgical skills. *Am.J.Obstet.Gynecol., 197,* 546-4.

Nussbaum, M. S. (2002). Surgical endoscopy training is integral to general surgery residency and should be integrated into residency and fellowships abandoned. *Semin.Laparosc.Surg., 9,* 212-215.

NVOG-HOOG 2005: http://www.medischevervolgopleidingen.nl//content/documenten/specialisme/gynaecologie/nvog_hoog.pdf

Omar, A. M., Wade, N. J., Brown, S. I., & Cuschieri, A. (2005). Assessing the benefits of “gaze-down” display location in complex tasks. *Surg.Endosc., 19,* 105-108.

Palter, V. N., Orzech, N., Aggarwal, R., Okrainec, A., & Grantcharov, T. P. (2010). Resident perceptions of advanced laparoscopic skills training. *Surg.Endosc.*.

Pandey, V. A., Wolfe, J. H., Liapis, C. D., & Bergqvist, D. (2006). The examination assessment of technical competence in vascular surgery. *Br.J.Surg., 93,* 1132-1138.

Park, A., Witzke, D., & Donnelly, M. (2002a). Ongoing deficits in resident training for minimally invasive surgery. *J.Gastrointest.Surg., 6,* 501-507.

Park, A. & Witzke, D. B. (2002b). Training and educational approaches to minimally invasive surgery: state of the art. *Semin.Laparosc.Surg., 9,* 198-205.

Pearson, A. M., Gallagher, A. G., Rosser, J. C., & Satava, R. M. (2002). Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying. *Surg.Endosc., 16,* 130-137.

Perkins, N., Starkes, J. L., Lee, T. D., & Hutchison, C. (2002). Learning to use minimal access surgical instruments and 2-dimensional remote visual feedback: how difficult is the task for novices? *Adv.Health Sci.Educ.Theory.Pract., 7,* 117-131.

Porte, M. C., Xeroulis, G., Reznick, R. K., & Dubrowski, A. (2007). Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *Am.J.Surg., 193,* 105-110.

Purkayastha, S., Aziz, O., Athanasiou, T., Paraskevas, P., & Darzi, A. (2004). Does laparoscopic surgery offer adequate clearance in rectal cancer?--A discussion. *Int.J.Surg., 2,* 103-106.

Rattner, D. W., Apelgren, K. N., & Eubanks, W. S. (2001). The need for training opportunities in advanced laparoscopic surgery. *Surg.Endosc., 15,* 1066-1070.

RCOG 2009: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/Ed-Core-OSATS.pdf

Reznick, R., Regehr, G., MacRae, H., Martin, J., & McCulloch, W. (1997). Testing technical skill via an innovative “bench station” examination. *Am.J.Surg., 173,* 226-230.

Reznick, R. K. & MacRae, H. (2006). Teaching surgical skills--changes in the wind. *N.Engl.J.Med., 355,* 2664-2669.

Rosser, J. C., Jr., Lynch, P. J., Cuddihy, L., Gentile, D. A., Klonsky, J., & Merrell, R. (2007). The impact of video games on training surgeons in the 21st century. *Arch.Surg., 142,* 181-186.

Rosser, J. C., Rosser, L. E., & Savalgi, R. S. (1997). Skill acquisition and assessment for laparoscopic surgery. *Arch.Surg., 132,* 200-204.

Schijven, M. & Jakimowicz, J. (2003). Virtual reality surgical laparoscopic simulators. *Surg. Endosc., 17,* 1943-1950.

Schijven, M. P. & Bemelman, W. A. (2011). Problems and pitfalls in modern competency-based laparoscopic training. *Surg.Endosc., 25,* 2159-2163.

Schijven, M. P., Berlage, J. T., & Jakimowicz, J. J. (2004). Minimal-access surgery training in the Netherlands: a survey among residents-in-training for general surgery. *Surg.Endosc., 18,* 1805-1814.

Schijven, M. P., Jakimowicz, J. J., Broeders, I. A., & Tseng, L. N. (2005). The Eindhoven laparoscopic cholecystectomy training course--improving operating room performance using virtual reality training: results from the first E.A.E.S. accredited virtual reality trainings curriculum. *Surg.Endosc., 19,* 1220-1226.

Schijven, M. P., Schout, B. M., Dolmans, V. E., Hendrikx, A. J., Broeders, I. A., & Borel Rinkes, I. H. (2008). Perceptions of surgical specialists in general surgery, orthopaedic surgery, urology and gynaecology on teaching endoscopic surgery in The Netherlands. *Surg.Endosc., 22,* 472-482.

Schreuder, H.W., van den Berg, C.B., Hazebroek, E.J., Verheijen, R.H., Schijven, M.P. (2011).

Laparoscopic skills training using inexpensive box trainers: which exercises to choose when constructing a validated training course. *BJOG., 118,* 1576-1584.

Scott, D. J., Bergen, P. C., Rege, R. V., Laycock, R., Tesfay, S. T., Valentine, R. J. et al. (2000). Laparoscopic training on bench models: better and more cost effective than operating room experience? *J.Am.Coll.Surg., 191,* 272-283.

Scott, D. J., Young, W. N., Tesfay, S. T., Frawley, W. H., Rege, R. V., & Jones, D. B. (2001). Laparoscopic skills training. *Am.J.Surg., 182,* 137-142.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K. et al. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann.Surg., 236,* 458-463.

Sharma, D., Shaban, A., Riddell, A., Kalsi, V., Arya, M., & Grange, P. (2009). Video-games station or minimally invasive skills training station? *BJU. Int., 104,* 159-160.

Shay, B. F., Thomas, R., & Monga, M. (2002). Urology practice patterns after residency training in laparoscopy. *J.Endourol., 16,* 251-256.

Sinha, P., Hogle, N. J., & Fowler, D. L. (2008). Do the laparoscopic skills of trainees deteriorate over time? *Surg.Endosc., 22,* 2018-2025.

Smith, S. G., Torkington, J., Brown, T. J., Taffinder, N. J., & Darzi, A. (2002). Motion analysis. *Surg. Endosc., 16,* 640-645.

Smith, W. D., Berguer, R., & Nguyen, N. T. (2005). Monitor height affects surgeons' stress level and performance on minimally invasive surgery tasks. *Stud.Health Technol.Inform., 111,* 498-501.

Southern Surgeons Club (1991). A prospective analysis of 1518 laparoscopic cholecystectomies. The Southern Surgeons Club. *N.Engl.J.Med., 324,* 1073-1078.

Stassen, L. P., Bemelman, W. A., & Meijerink, J. (2010). Risks of minimally invasive surgery underestimated: a report of the Dutch Health Care Inspectorate. *Surg.Endosc., 24,* 495-498.

Stefanidis, D., Acker, C. E., Swiderski, D., Heniford, B. T., & Greene, F. L. (2008). Challenges during the implementation of a laparoscopic skills curriculum in a busy general surgery residency program. *J.Surg.Educ., 65,* 4-7.

Stefanidis, D. & Heniford, B. T. (2009a). The formula for a successful laparoscopic skills curriculum. *Arch.Surg., 144,* 77-82.

Stefanidis, D., Hope, W. W., Korndorffer, J. R., Jr., Markley, S., & Scott, D. J. (2010). Initial laparoscopic basic skills training shortens the learning curve of laparoscopic suturing and is cost-effective. *J.Am.Coll.Surg., 210,* 436-440.

Stefanidis, D., Korndorffer, J. R., Jr., Black, F. W., Dunne, J. B., Sierra, R., Touchard, C. L. et al. (2006a). Psychomotor testing predicts rate of skill acquisition for proficiency-based laparoscopic skills training. *Surgery, 140,* 252-262.

Stefanidis, D., Korndorffer, J. R., Jr., Markley, S., Sierra, R., & Scott, D. J. (2006b). Proficiency maintenance: impact of ongoing simulator training on laparoscopic skill retention. *J.Am. Coll.Surg., 202,* 599-603.

Stefanidis, D., Korndorffer, J. R., Jr., Sierra, R., Touchard, C., Dunne, J. B., & Scott, D. J. (2005). Skill retention following proficiency-based laparoscopic simulator training. *Surgery, 138,* 165-170.

Stefanidis, D., Scott, D. J., & Korndorffer, J. R., Jr. (2009b). Do metrics matter? Time versus motion tracking for performance assessment of proficiency-based laparoscopic skills training. *Simul.Healthc., 4,* 104-108.

Strom, P., Hedman, L., Sarna, L., Kjellin, A., Wredmark, T., & Fellander-Tsai, L. (2006). Early exposure to haptic feedback enhances performance in surgical simulator training: a prospective randomized crossover study in surgical residents. *Surg.Endosc., 20,* 1383-1388.

Sweet, R. M., Hananel, D., & Lawrenz, F. (2010). A unified approach to validation, reliability, and education study design for surgical technical skills training. *Arch.Surg., 145,* 197-201.

Swift, S. E. & Carter, J. F. (2006). Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. *Am.J.Obstet. Gynecol., 195,* 617-621.

Torkington, J., Smith, S. G., Rees, B., & Darzi, A. (2001a). The role of the basic surgical skills course in the acquisition and retention of laparoscopic skill. *Surg.Endosc., 15,* 1071-1075.

Torkington, J., Smith, S. G., Rees, B. I., & Darzi, A. (2001b). Skill transfer from virtual reality to a real laparoscopic task. *Surg.Endosc., 15,* 1076-1079.

Trimbos J.B. (2007). *Basics of Surgery, tools, techniques, attitude and expertise*.

van der Meijden, O. A. & Schijven, M. P. (2009). The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surg. Endosc., 23,* 1180-1190.

van Dongen, K. W., van der Wal, W. A., Rinkes, I. H., Schijven, M. P., & Broeders, I. A. (2008). Virtual reality training for endoscopic surgery: voluntary or obligatory? *Surg.Endosc., 22,* 664-667.

van Hove, P. D., Tuijthof, G. J., Verdaasdonk, E. G., Stassen, L. P., & Dankelman, J. (2010).

Objective assessment of technical surgical skills. *Br.J.Surg., 97,* 972-987.

Van Sickle, K. R., McClusky, D. A., III, Gallagher, A. G., & Smith, C. D. (2005). Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg.Endosc., 19,* 1227-1231.

Van Herzeele, I., O'Donoghue, K. G., Aggarwal, R., Vermassen, F., Darzi, A., & Cheshire, N. J. (2010). Visuospatial and psychomotor aptitude predicts endovascular performance of inexperienced individuals on a virtual reality simulator. *J.Vasc.Surg., 51,* 1035-1042.

Vassiliou, M. C., Feldman, L. S., Andrew, C. G., Bergman, S., Leffondre, K., Stanbridge, D. et al. (2005). A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am.J.Surg., 190,* 107-113.

Verdaasdonk, E. G., Stassen, L. P., Schijven, M. P., & Dankelman, J. (2007a). Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. *Surg.Endosc., 21,* 1406-1412.

Verdaasdonk, E. G., Stassen, L. P., van Wijk, R. P., & Dankelman, J. (2007b). The influence of different training schedules on the learning of psychomotor skills for endoscopic surgery. *Surg.Endosc., 21,* 214-219.

Vossen, C., Van Ballaer, P., Shaw, R. W., & Koninckx, P. R. (1997). Effect of training on endoscopic intracorporeal knot tying. *Hum.Reprod., 12,* 2658-2663.

Westebring-van der Putten EP, Goossens, R. H., Jakimowicz, J. J., & Dankelman, J. (2008). Haptics in minimally invasive surgery--a review. *Minim.Invasive.Ther.Allied Technol., 17,* 3-16.

Zehetner, J., Kaltenbacher, A., Wayand, W., & Shamiyeh, A. (2006). Screen height as an ergonomic factor in laparoscopic surgery. *Surg. Endosc., 20,* 139-141.

AUTHOR AFFILIATIONS

From the department of Medical statistics and Bioinformation,
Leiden University Medical Center, Leiden, the Netherlands
Ron Wolterbeek, Saskia le Cessie

From the department of Medical Decision making,
Leiden University Medical Center, Leiden, the Netherlands
Anne Stiggelbout

From the department of Gynaecologie,
Leiden University Medical Center, Leiden, the Netherlands
Wendela Kolkman, Baptist Trimbos, Frank Willem Jansen

From the department of Gynaecology
University Medical Center Utrecht, Utrecht, The Netherlands
Henk Schreuder

From the department of BioMechanical Engineering,
Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft
Magda Chmarra, John van den Dobbelsteen, Jenny Dankelman

Former medical students, currently busy with their careers in medical microbiology, ophthalmology and neurosurgery.
Liz Terveer, Matthijs van der Put, Navid Hossein pour Khaledian

ABOUT THE AUTHOR

Ellen Hiemstra was born on September 23rd 1979 in Zwolle. She attended the Gymnasium Celeanum in Zwolle, which she finished Cum Laude in 1997. She moved to Delft to study Civil Engineering at the Delft University of Technology. After obtaining her Propadeuse and meeting her future partner Rienk Krol, she was admitted to study Medicine at the University of Utrecht which she finished in 2004. She started to work as a resident at the department of Gynaecology of the Leiden University Medical Center (LUMC) and at the department of Obstetrics and Gynaecology of the Haga Hospital in the Hague. In 2005 she was selected for the official residency training program in Obstetrics and Gynaecology. After one year of training at the Haga Hospital (Dr. E.J. van Rijssel), she decided to combine her passion for medicine with her previous passion for technology. She attended a PhD program (2007-2009) at the department of Minimally invasive surgery in gynaecology at the LUMC (Prof. Dr. F.W. Jansen) in collaboration with the Delft University of Technology (Prof. J. Dankelman), resulting in this thesis. In January 2010 she continued her residency training at the Haga Hospital (Dr. B.W. Hellebrekers) and currently at the LUMC (Prof. Dr. J.M. van Lith). In the continuum of the circle of life, Malou (the daughter of Rienk and Ellen) starts exploring the realm of technology at home with plugs and electrical sockets.

PUBLICATIONS

Related to this thesis

Hiemstra E, Kolkman W, Jansen FW. Skills training in minimally invasive surgery in Dutch obstetrics and gynecology residency curriculum. Gyn Surg 2008; 5: 321-5

Hiemstra E, Kolkman W, Put van de MAJ, Jansen FW. Retention of basic laparoscopic skills after a structured training program. Gyn Surg 2009; 6 229-35

Hiemstra E, Kolkman W, Wolterbeek R, Trimbos JBMZ, Jansen FW. The Value of an Objective Assessment Tool in the Operating Room. Can J Surg, 2011; 54: 116-22

Hiemstra E, Kolkman W, le Cessie S, Jansen FW. Are Minimally Invasive Procedures harder to acquire than Conventional Surgical Procedures? Gyn Obstet Invest. 2011; 71: 268-73.

Hiemstra E, Terveer EM, Chmarra MK, Dankelman J, Jansen FW. Virtual Reality in Laparoscopic Skills Training: Is Haptic Feedback replaceable? Minimally Invasive Therapy & Allied Techniques, 2011; 20: 79-84

Hiemstra E, Jansen FW. Richtlijn OSATS. Tijdschrift voor Medisch Onderwijs. 2010. 29; 5: 135-41.

Hiemstra E, Chmarra MK, Dankelman J, Jansen FW. Intracorporeal suturing: economy of movements in a box trainer model. JMIG, 2011; 18: 494-9

Jansen FW, **Hiemstra E**. Comment on *Laparoscopic skills training using inexpensive box trainers: which exercises to choose when constructing a validated training course*. BJOG, 2011; accepted

Hiemstra E, Hossein pour Khaledian N, Trimbos JBMZ, Jansen FW. Intraoperative assessment of resident's surgical skills: benefits and pitfalls. *Submitted*

Hiemstra E, Schreuder HWR, Stiggelbout AM, Jansen FW. Grading surgical skills curricula and training facilities for minimally invasive surgery. *Submitted*

Hiemstra E, Hossein pour Khaledian N, Dobbelsteen JJ, Dankman J, Jansen FW. Optimizing laparoscopic skills training: Does a fixed Camera Compromise Depth Perception? *Submitted*

Other publications

Hiemstra E, Weijenborg PT, Jansen FW. Management of chronic pelvic pain additional to tubal sterilization. J Psychosom Obstet Gynaecol. 2008 :29 (3) 153 -6

Schaafsma BE, **Hiemstra E**, Dankelman J, Jansen FW. Feedback in laparoscopic skills acquisistion: an observational study during a basic skills training course. Gynecol Surg. 2009; 6: 339–43

Kroon de CD, **Hiemstra E**, Trimbos JBMZ, Jansen FW. Power Doppler area in the diagnosis of endometrial cancer. Int J Cancer. 2010; 20: 1160-5

Okken LM, Chmarra MK, **Hiemstra E**, Dankelman J, Jansen FW. Assessment of joystick and wrist control in hand-held articulated laparoscopic prototypes. *Submitted*

&

Hiemstra E, Schoot D. Hoofdstuk 10 Skills oefeningen. Hysteroscopisc Chirurgie: de basis. Jansen FW, Trimbos-Kemper T. Editor Laurier, Noordwijk, 2008. ISBN: 9789081289719

Hiemstra E, Terveer EM, Dankelman J, Jansen FW. Comment on Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. Website Br J Surg 2008.

Rodrigues SP, de Burlet KJ, **Hiemstra E**, Twijnstra ARH, van Zwet EW, Trimbos-Kemper TCM, Jansen FW, Dynamics of b-hCG in ectopic pregnancy, submitted

&

DANKWOORD

Mijn dank gaat uit naar de vele AIOS die voor dit proefschrift bereid bleken hun beoordelingen op de operatiekamer met mij te delen. Jullie werkten je in het zweet om een laparoscopische hechting te leggen, en lieten soms zelfs speekselafnames op de operatiekamer toe zodat ik het stressniveau kon meten.

Daarnaast wil ik de alle studenten bedanken voor hun telkens weer enthousiaste deelname aan experimenten in de oefenbox en op VR trainer. Door jullie enthousiasme vond ik mijn eigen onderzoek vaak nog, en heel soms weer, leuker.

Ook ben ik de gynaecologen dankbaar voor al die keren dat zij het weefselgevoel en de instrumentenkennis van AIOS in maat en getal wilden uitdrukken op een OSATS. Jullie kritisch meedenken over de OSATS en over opleiden in het algemeen is heel waardevol voor mij geweest.

Zowel binnen het LUMC, als op de TU Delft zijn er vele afdelingen waar de mensen mij steeds weer een warm welkom gaven met zowel hulp als gezelligheid. Ik denk aan alle helpende handen en luisterende oren op het stafsecretariaat op K6, aan de 5^e verdieping van het onderzoeksgebouw voor statistische ondersteuning met chocotoffs, en het skillslab waarin kipfilets, paprika's en strijkkralen een tweede leven kregen. Natuurlijk vergeet ik het Biomechanical Engineering laboratorium niet met uitvindingen variërend van ruggenpriksimulatoren tot operatieopstellingen. Hier bleken simpele prikstokken van de plantsoenendienst te kunnen worden omgebouwd tot laparoscopisch instrumentarium - dit alles voor de wetenschap.

Allerbeste J7 vrienden, jullie maakten mijn promotietraject tot een feest! Goede herinneringen heb ik aan ons experiment waarbij we randomiseerden tussen een dag zeilen en een dag artikelen lezen, met als uitkomst wetenschappelijke vordering. Fijn hoe onze serieuze discussies steeds afgewisseld werden door relativerende grappen!

Familie en vrienden, dank voor jullie aanhoudende interesse in (en aan het einde soms bewust maar even niet vragen naar) mijn proefschrift. In het bijzonder dank aan mijn mijn creatieve vriendin Marjon die zich op de omslag heeft gestort, aan mijn neef Hylke die vanuit Amerika enkele kromme engelse zinnen weer lopende heeft gemaakt en natuurlijk mijn lieve ouders en broer, voor alles.

Stoere, lieve Rienk, elke keer als ik in proefschriften verontschuldigingen aan partners las voor de vele jaren verwaarlozing, dacht ik dat wij het samen nog best goed voor elkaar hadden. Toch is het groot feest nu dit ei, zeker met jouw steun, is gelegd. Het wordt nu nog leuker!

Ellen