



PhD thesis

Ebbe Thinggaard

Take-Home Training in Laparoscopy

Academic advisors:

Ismail Gögenur

Lars Konge

This thesis has been submitted to the Graduate School of the Faculty of Health and Medical Sciences, University of Copenhagen.

09/08/2016

Author: Ebbe Thinggaard, MD
Title: Take-Home Training in Laparoscopy
Danish Title: Hjemmetræning i kikkertkirurgi
Date of Submission: 9th of August 2016.

Academic Advisors:

Professor Ismail Gögenur, MD, DMSc
Department of Surgery
Zealand University Hospital
University of Copenhagen

Associate Professor Lars Konge, MD, PhD
Copenhagen Academy for Medical Education and Simulation
Capital Region, Centre of Human Resources
University of Copenhagen

Assessment Committee:

Professor Jacob Rosenberg, MD, DMSc
Department of Surgery
Herlev University Hospital
University of Copenhagen

Professor Lars Lund, MD, DMSc
Department of Urology
Odense University Hospital
University of Southern Denmark

Professor Teodor Grantcharov, MD, PhD
Department of Surgery, St. Michael's Hospital
University of Toronto
Canada Research Chair in Simulation and Surgical Safety
Keenan Research Centre of the Li Ka Shing Knowledge Institute
Minimally Invasive Surgery

ACKNOWLEDGEMENTS

First of all I would like to thank my two academic advisors, Ismail Gögenur and Lars Konge. I got to know Ismail when I was medical student and he volunteered to teach on our student-organised laparoscopy course. From then on, there was no doubt that if I were to do a PhD, I would want Ismail as my supervisor. Thank you for your help, patience and guidance into the world of academia. I got to know Lars when we started planning the PhD, and what better time to start a PhD than when the Simulation centre was starting to take off. I thank you for your thorough, direct and honest feedback, and for helping me begin a career in medical education.

Apart from my academic advisors, Flemming Bjerrum and Jeanett Strandbygaard have been part of the research group throughout the PhD. Thank you for creating the foundation from which this PhD has emerged and thank you for your collaboration, I wish you all the best in your future careers.

Throughout the PhD I have had the good fortune of working with a number of people; Jakob Kleif, Matt Ritter and Lene Spanger, thank you for your help. I owe Lene much gratitude for her help in conducting the qualitative parts of my research and without your help I would probably still be trying to figure out what to do with my interview transcripts.

I would also like to thank the two research groups I have been a part of these last couple of years. Thank you to my colleagues at the Centre of Surgical Science and my colleagues at the Copenhagen Academy for Medical Education and Simulation. Two people have been particularly helpful in making this PhD possible. Thank you Tove and Leizl for your help in overcoming all the practical hurdles along the way. Without such colleagues the journey would have been much, much harder.

A special thanks to my family and my wife Sohur. I know the PhD years have been bumpy and I appreciate your support in getting through the rough parts and being there to enjoy the good parts as well.

Ebbe Thinggaard, Copenhagen, August 2016

Table of Contents

PAPERS INCLUDED IN THE THESIS	6
SUMMARY	7
DANSK RESUMÉ.....	9
1. INTRODUCTION	12
2. BACKGROUND.....	13
Development of Operative Laparoscopy	13
Training of Laparoscopic Techniques	13
Take-Home Training	14
3. THEORETICAL CONSIDERATIONS	16
Proficiency-Based Learning.....	16
Distributed Practice	16
Self-Regulation in Learning.....	17
4. METHODOLOGICAL CONSIDERATIONS	18
Setting.....	18
A Mixed-Methods Approach	19
Validity	20
Standard Setting	21
Focus Groups and Individual Interviews	22
Statistical Considerations	23
Ethical Considerations	23
Overall Purpose of the Thesis	24
Aims.....	24

Hypotheses	24
5. PRESENTATION OF THE INCLUDED PAPERS	26
6. DISCUSSION.....	42
7. CONCLUSIONS.....	47
8. IMPLICATIONS AND FUTURE RESEARCH	47
9. REFERENCES.....	49
10. PAPERS	59

PAPERS INCLUDED IN THE THESIS

The Thesis is based on the following papers:

Paper 1

Thinggaard E, Kleif J, Bjerrum F, Strandbygaard J, Gögenur I, Ritter E M, Konge L. Off-site training of laparoscopic skills, a scoping review using a thematic analysis. *Surgical Endoscopy* 2016;11:1-9.

Paper 2

Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *British Journal of Surgery* 2015;102:1106–1113.

Paper 3

Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Ensuring competency of novice laparoscopic surgeons – Exploring standard setting methods and their consequences. *Journal of Surgical Education* 2016.

Paper 4

Thinggaard E, Bjerrum F, Strandbygaard J, Konge L, Gögenur I. Take-home training in laparoscopy facilitates distributed training, a randomized trial. *Surgery* 2016; Submitted.

Paper 5

Thinggaard E, Konge L, Bjerrum F, Strandbygaard J, Gögenur I, Spanager L. Take-home training in a simulation-based laparoscopy course. *Surgical Endoscopy* 2016; In Print.

SUMMARY

When laparoscopy was first introduced, skills were primarily taught using the apprenticeship model. A limitation of this method when compared to open surgery, was that it requires more time to practise and more frequent learning opportunities in clinical practice. The unique set of skills required in laparoscopy highlighted the need for new training methods that reduce the need for supervision and do not put the patient at risk. Simulation training was developed to meet this need. The overall purpose of this thesis was to explore simulation-based laparoscopic training at home. The thesis consists of five papers: a review, a validation study, a study of methodology, a randomised controlled trial and a mixed-methods study. Our aims were to review the current knowledge on training off-site, to develop and explore validity for a training and assessment system, to investigate the effect of take-home training in a simulation-based laparoscopic training programme, and to explore the use of take-home training.

The first paper in this thesis is a scoping review. The aim of the review was to explore the current knowledge on off-site laparoscopic skills training. We found that off-site training was feasible but that changes were required in order for it to become an effective method of training. Furthermore, the selected instructional design varied and training programmes were designed using a variety of educational theories. Based on our findings, we recommended that courses and training curricula should follow established education theories such as proficiency-based learning and deliberate practice. Principles of directed self-regulated learning could be used to improve off-site laparoscopic training programmes.

In the second study, we set out to develop and explore validity evidence of the TABLT test. The TABLT test was developed for basic laparoscopic skills training in a cross-specialty curriculum. We found validity evidence to support the TABLT test as a summative test in a basic laparoscopic training programme. We also established a credible pass/fail level using the contrasting groups method. We concluded that the TABTL test could be used to assess novice laparoscopic trainees across different specialties and help trainees acquire basic laparoscopic competencies prior to supervised surgery.

In the third study, we aimed to explore the consequences of the choice of standard setting method and whether there is a difference in terms of how high a score experienced and novice laparoscopic surgeons expect that novices should achieve during training. We used three different standard setting methods and found that pass/fail levels vary depending on the choice of standard setting method. We also asked experienced and novice laparoscopic surgeons how high a score they expected a novice laparoscopic surgeon should achieve on a test during training. We

found a significant difference, with experienced surgeons setting a lower pass/fail level. We concluded that an established standard setting method supported by evidence should be used when setting a pass/fail level.

In the first and second papers of this thesis, we found that off-site training is feasible and explored validity for the TABLT test. We used this knowledge in the fourth study to design a randomised controlled trial. The aim of the trial was to investigate the effect of take-home training in a simulation-based laparoscopic course. We hypothesised that training at home could help trainees plan their training according to their own schedule and thereby increase the effect of training. We found that participants had a distributed training pattern; they trained more frequently and in shorter sessions. We also found that participants were able to rate their own performance during unsupervised training and that self-rating was reliable.

The fifth and final study of the thesis was a mixed-methods study that aimed to explore the use of take-home training. To meet this aim, we recruited participants from the intervention arm in our randomised controlled trial. All participants had access to the simulation centre and were given a portable trainer to train on at home. Participants were asked to use a logbook during training. At the end of the course, they were invited to take part in focus group interviews and individual interviews. Based on data from logbooks, a descriptive statistical analysis was conducted and data from interviews were analysed using a content analysis. We found that participants took an individualised approach to training when training at home. They structured their training according to their needs and external requirements. We concluded that mandatory training requirements and testing help determine when and how much participants train. We also found that self-rating can guide unsupervised training by giving clear goals to be reached during training.

From the papers included in the thesis, we found that the literature describes training at home as a feasible method of acquiring laparoscopic skills. Nonetheless, changes to current training programmes are needed in order to make this method effective. We then developed and explored validity evidence for the TABLT test. We also established a reasonable pass/fail level and went on to explore the immediate consequences of the pass/fail level. Using our knowledge from the review, we conducted a randomised controlled trial and a mixed-method study. Based on these studies we found that training at home allows for distributed learning, that self-rating guides unsupervised training, and that mandatory training requirements and testing strongly influence training patterns. Access to training, guidance during training, and mandatory training requirements will make take-home training not just feasible but also effective.

DANSK RESUMÉ

Da laparoskopien blev introduceret, brugte man principperne om mesterlære, til at undervise i de tekniske færdigheder. En begrænsning ved laparoskopi i forhold til åben kirurgi, er at man skal bruge længere tid på at øve sig. Det kræver at man besidder unikke færdigheder, for at udføre laparoskopiske indgreb og derfor opstod der et behov for nye undervisningsmetoder. Metoder der kunne nedsætte brugen af superviseret oplæring samt øge patientsikkerheden. Simulationstræning blev udviklet for at imødekomme dette behov. Det overordnede mål med denne afhandling er at undersøge simulationstræning i kikkertkirurgi derhjemme. Afhandlingen består af fem studier, et review, et valideringsstudie, et metode studie, et randomiseret studie og et "mixed-methods" studie. Formålet med studierne var at klarlægge den nuværende viden om træning i kikkertkirurgi derhjemme, at udvikle og undersøge validitetsevidens for en test i kikkertkirurgi, at undersøge konsekvenserne af den metode man bruger når man sætter beståelsesniveauet på en test, at undersøge effekten af at træne derhjemme samt at undersøge hvordan kursister træner derhjemme.

Det første studie i afhandlingen er et "scoping review". Formålet var at klarlægge den eksisterende viden om træning i kikkertkirurgi derhjemme. Vi fandt ud at træning derhjemme er muligt, men at det er nødvendigt med forbedringer, før det kan blive en effektiv metode. Endvidere viste det sig, at der bliver brugt forskellige uddannelsesteorier til at udvikle kurser med hjemmetræning. På baggrund af vores fund, anbefalede vi at man brugte veletablerede uddannelsesteorier, når man udvikler kurser med hjemmetræning.

I det andet studie udviklede vi TABLT testen og undersøgte validitetsevidens. TABLT testen blev udviklet til tværspeciale kurser i kikkertkirurgi, hvor læger fra forskellige specialer træner sammen. Vi fandt ud af, at der er evidens der understøtter validiteten af TABLT testen. Endvidere fastlagde vi en beståelsesgrænse, der var fair. Vi konkluderede, at man ville kunne bruge TABLT testen, til at lære basale kikkertkirurgiske færdigheder, før man udfører superviserede laparoskopiske indgreb på patienter.

I det tredje studie undersøgte vi konsekvenserne af den metode, man bruger til at fastsætte beståelsesniveauer på. For at undersøge konsekvenserne af metodevalg brugte vi tre forskellige metoder til at sætte beståelsesniveauer på. Vi fandt ud af, at beståelsesniveauet afhang af hvilken metode man brugte. Endvidere spurgte vi erfarne og uerfarne, hvor højt de mente, at uerfarne kikkertkirurger ville kunne score på en test i løbet af et træningsforløb. Vi fandt en signifikant forskel, da erfarne satte et lavere beståelsesniveau end uerfarne. Vi konkluderede, at et beståelsesniveau skal sættes ved at bruge en metode der støttes af evidens.

I første og andet studie fandt vi ud af at træning derhjemme var muligt og at der er validitetsevidens for TABLT testen. Vi brugte denne viden, da vi designede det fjerde studie. Her udførte vi et randomiseret forsøg, hvor vores hypotese var, at hjemmetræning ville give læger mulighed, for at planlægge deres træning efter deres behov og dermed øge effekten af at træne. Vi fandt ud af, at kursister der trænede derhjemme, trænede mere fordelt, dvs. at de trænede oftere, men i kortere intervaller. Vi fandt også ud af, at læger er i stand til at vurdere deres egen test.

I det femte og sidste studie, designede vi et "mixed-methods" forsøg. Formålet med forsøget var, at undersøge hvordan læger træner derhjemme. Deltagere fra det randomiserede forsøg indgik som forsøgspersoner og blev bedt om at bruge logbøger til at dokumentere deres træning. Vi inviterede dem også til at deltage i fokusgruppe interviews og individuelle interviews. Data fra logbøgerne brugte vi til at lave en kvantitativ deskriptiv statistisk analyse og data fra interviews brugte vi til at lave en kvalitativ indholdsanalyse. På baggrund af disse analyser, fandt vi ud af, at forsøgspersonerne brugte en individuel tilgang til at træne derhjemme. De strukturerede deres træning, i forhold til deres behov og eksterne krav. Vi konkluderede, at krav om træning i form af en test, er bestemmende for hvordan og hvor meget man træner derhjemme. Endvidere fandt vi ud af, at selvvrurderet testning gav forsøgsdeltagere et klart mål i løbet af deres træning.

Fra de studier der indgår i denne afhandling har vi kunnet konkludere følgende: Litteraturen beskriver at træning derhjemme er muligt, men at der er behov for ændringer i træningsprogrammerne hvis de også skal være effektive. Vi har endvidere udviklet en test i basale kikkertkirurgiske færdigheder og udforsket validitetsevidens for denne test. Vi har fastsat et fair beståelsesniveau og derefter undersøgt konsekvenserne af metodevalg, når man fastsætter et beståelsesniveau. Baseret på den viden, vi havde opnået i de første studier udførte vi herefter et randomiseret forsøg og siden et "mixed-methods" forsøg. På baggrund af disse forsøg, fandt vi ud af, at træning derhjemme resulterer i et mere opdelt træningsmønster, at læger er i stand til at vurdere deres egne tests, at disse selvvrurderinger kan bruges til at støtte deres træning og at træningskrav i form af obligatoriske testning er bestemmende for, hvordan og hvor meget man træner derhjemme. Ved at implementere en bedre adgang til træning, støtte i løbet af træningen og et krav om træning, kan man ikke bare muliggøre hjemmetræning, men også gøre det til en effektiv træningsmetode.

ABBREVIATIONS

VRS	Virtual reality simulator
BT	Boxtrainer
MIS	Minimal invasive surgery
FLS	Fundamentals of laparoscopic surgery
TABLT	Training and assessment of basic laparoscopic techniques
SRL	Self-regulated learning
SDL	Self-directed learning
DSRL	Directed self-regulated learning
PBL	Problem-based learning
CBME	Competency-based medical education
ANOVA	Analysis of variance
ICC	Intraclass correlation coefficient

1. INTRODUCTION

Minimally invasive surgery, and laparoscopy in particular, have revolutionised surgical care (1). The implementation of operative laparoscopy has reduced the duration of hospital stay and the convalescence period and has helped improve patient outcomes and enhance recovery after surgery (2, 3). Laparoscopy has gone through a series of developmental phases and has become the primary choice as a surgical technique (4). Laparoscopy is what surgical trainees across the different specialties are trained to do when performing surgery. Although the laparoscopic technique has many benefits, implementation was a challenge. Laparoscopic skills are very different to those used in open surgery and require specific training (5). The need for more training was recognised in the early phases of the development of the laparoscopic technique (6, 7). In the United States, training and passing a test is now a requirement in order to be certified as a general surgeon to be proficient in laparoscopy (8).

Originally, laparoscopic skills were primarily taught using the apprenticeship model. However, this method has certain limitations: it requires longer time to practice and more learning opportunities in clinical practice. The unique set of skills required in laparoscopy highlighted the need for new training methods, so simulation training was developed. During the last few decades, it has been firmly established that laparoscopic skills can be acquired outside the OR using simulators (9). Training can be done on either virtual reality simulators (VRSs) or boxtrainers (BTs); both methods have been shown to be effective methods for providing laparoscopic skills training (10).

Despite the evidence for the use of simulation training, implementation has been unsystematic (11). Barriers to simulation training – including the need for time to train, the high price of simulation equipment, and the lack of access – have halted the implementation of simulation training in laparoscopic training programmes (12). Low accessibility in particular has been a constraint. Access to training can be improved by using mobile BTs, as this method is affordable, accessible and mobile (12). BTs allow trainees to train according to their own schedule. Furthermore, BTs have the potential to improve laparoscopic training when implemented in a simulation-based laparoscopic training programme. Unfortunately, little is known about the use of BTs in off-site training. No review has been done to explore this, and little research has been conducted into the effects of off-site training.

2. BACKGROUND

Development of Operative Laparoscopy

The move towards minimally invasive surgery can be traced back as far as Hippocrates in ancient Greece (13) and Abu Al Qasim Al-Zahrawi in medieval Spain (1). Both Hippocrates and Al-Zahrawi developed speculums to allow access to body cavities in order to alleviate symptoms and treat diseases. Minimally invasive surgery has gone through a series of developmental phases since then. What we have come to know as operative laparoscopy today was primarily developed in the 1970s, when the technique was explored and many new devices were invented (13). The breakthrough in laparoscopy came when it was used for cholecystectomies (1). In 1987, Mouret was the first to perform a laparoscopic cholecystectomy, and soon others followed. There were initially challenges with the large-scale implementation. The initial increase in bile duct injuries illustrated the problems associated with implementing a new technique at such a quick pace (7). Despite the increase in bile duct injuries, laparoscopic cholecystectomy was implemented for more widespread use (1). The benefits of laparoscopy include the reduction of the duration of hospital stay and convalescence period (2, 3). Laparoscopy has since become the preferred technique for intra-abdominal surgery (4). Although the technique was initially used for benign surgery, it has also found a use in oncological procedures, where the same benefits have been demonstrated without compromising the oncological outcome (14).

Training of Laparoscopic Techniques

Laparoscopy requires a very different technique compared to open surgery. When operating using the laparoscopic technique, the surgeon uses a set of very specific skills (5). The surgeon will have to accommodate for the loss of depth perception and the limited range of movement and adjust to the use of long instruments fixed at skin level. However, the necessary skills can be acquired through training. Training is needed to become proficient in laparoscopy and a new surgical curriculum for laparoscopy was needed (6, 7). In the early days of laparoscopy, training was done in the same manner as for open surgical skills. Training was based on the apprenticeship model and primarily conducted during supervised surgery in the operating room (OR) (15). Nonetheless, the Hallstead approach, often referred to as ‘see one, do one, teach one’, was found to be insufficient for laparoscopic skills training. The Hallstead model required a long mentoring process, which was neither cost-effective nor compatible with the increased awareness of the potential risk to patients from having untrained surgeons performing laparoscopy (15). Simulation-based training was

suggested as a solution and BTs were developed (16). Even VRSs were proposed early on as a potential training method (17-19). Although research has firmly established that laparoscopy can be taught outside the OR using simulation training (9), it has not been implemented systematically (11). A body of evidence supports the argument that VRSs and BTs are effective methods of acquiring laparoscopic skills (9). Laparoscopic simulation-based training shortens operating time, increases intra-operative skills, and reduces the risk of intra-operative and post-operative complications (20-24). Although many VRSs and BTs exist, systematic training programmes are lacking and have only been implemented in a few countries (11). An example of this is the FLS, which is a requirement for becoming a general surgeon in the USA (8). However, the FLS is not accessible to most surgeons or to those most in need of laparoscopic simulation training. The tasks included in the FLS are advanced and are therefore inappropriate for novices. Requirements in training should correspond to the trainee's level of experience (25). There is a need for a new test that is appropriate, affordable and accessible to novices.

Testing and the pass/fail level of a test provide the minimum requirement for training in a proficiency-based training programme. Proficiency-based training has been recommended for simulation-based laparoscopy training and has been shown to be effective (26, 27). Training programmes using proficiency-based training rely on the use of testing and pass/fail levels. Standard-setting methods are used when setting a pass/fail level for a test. However, the pass/fail level of the test may vary considerably depending on the standard setting method (28). Despite this, few studies have explored the consequences of the choice of standard setting method in a proficiency-based laparoscopic training programme.

Take-Home Training

Proficiency-based laparoscopic training programmes have been developed for BTs (29). Training laparoscopic skills on BTs is an effective method of acquiring laparoscopic skills (30), and a variety of BTs have been developed for this purpose. These range from BTs that are similar to the OR setting where laparoscopic cameras are used, to simple trainers (31) and even open trainers to practice moving instruments (32). Some BTs include an eye-patch to remove stereovision (33) and other trainers use mirrors (34, 35). BTs have even been developed to address ergonomics (36). The development of BT designs has followed that of digital cameras. The early BTs used large home video cameras (16) and web-cameras were also used (37, 38). Both do-it-yourself (DIY) trainers (16, 39, 40) and commercially available ones have been suggested (41). Recently, tablet computers such as the iPad (31, 42) and even telephones (43) have been used in BT design. The on-going development of BTs have made them accessible, affordable and mobile (12).

Mobility in training can help overcome some of the barriers to simulation training. Times to train, access and the high price for simulation equipment are barriers to simulation training (12). Mobile BTs make training accessible to novice laparoscopic trainees as they can plan their training according to their own schedule. Despite this apparent advantage, no review has yet explored the literature regarding off-site training. BTs may not only provide trainees with the flexibility they need but could also prove beneficial when implemented in existing simulation-based training programmes. Currently, simulation-based training programmes primarily offer training at a simulation centre or skills lab. Additionally training at home could improve current training programmes and increase the use of laparoscopic simulation training in general.

3. THEORETICAL CONSIDERATIONS

Laparoscopic training programmes should be based on sound educational theories. Research in medical education must be placed in a theoretical framework in order to support evidence of findings (44). Below we outline the most prominent educational theories used in this thesis.

Proficiency-Based Learning

In the early 1900s the Flexner Report made it clear that competency-based medical education (CBME) was necessary to ensure high-quality training in medicine (45, 46). CBME has recently undergone a revival in post-graduate training and surgical skills training. Passing a competency-based test is now a requirement in the USA in order to obtain specialty registration as a general surgeon (8). CBME is being introduced as proficiency-based training in laparoscopic skills training (15, 27). Proficiency-based training is a further development of CBME, where the minimum requirement is that of proficiency rather than competency. Research has shown the effect of proficiency-based training on OR performance (21, 29). Proficiency-based training is also recommended for simulation training in laparoscopy (27). Proficiency-based training relies on the use of assessment and testing. Testing has a positive effect on retention of learned skills (47). Proficiency-based standard setting has been implemented using performance levels of experienced laparoscopic surgeons (48). However, when implementing tests it is important that they are supported by evidence of validity (49). Testing should be an integrated part of the training curriculum in which learning objectives, content, assessment and use of skills in the OR are aligned. Proficiency-based training and testing are discussed in the papers included in this thesis.

Distributed Practice

Distributed practice, or spaced repetition as it is sometimes referred to, is an effective educational approach. Distributed practice is superior to massed practice as learners can divide their learning into manageable parts (50). This is also referred to as the spacing effect. Distributing learning experiences across a number of days is effective and may also apply in motor skills training (51). It is recommended for gaining knowledge and motor skills (52, 53). The effect has also been demonstrated in laparoscopic skills training (54, 55). Nonetheless, the optimal distribution of practice remains to be established (56). Distributed practice is also a part of deliberate practice, which is effective in procedural skills training (57). In this thesis, distributed practice is discussed in the first, fourth and fifth papers.

Self-Regulation in Learning

Self-regulation in learning relies on theories from both pedagogy and psychology. Two terms that are often used to describe self-regulation in learning are SRL and SDL. SRL is sometimes confused or used interchangeably with SDL, and these two terms have been used in various ways in the literature (58). SRL is when one considers the students' independence in training, whereas SDL is used when training promotes autonomous learning. A new theory in medical education is DSRL, which is based on theories of SRL. DSRL is a new approach to learning in which trainees regulate their own training within a framework provided by educators. Trainees thereby control a part of the training and are active participants in their own learning process. Faculty act as facilitators to guide the self-regulated learning by providing a framework in which learners operate (59). SRL is the basis for DSRL, where the focus is on understanding autonomous learning to help provide guidance during learning experiences (60). SRL is recommended in a recent systematic review for simulation-based training (61). One study has proven DSRL to be effective for retention of skills in a simulation training for lumbar puncture (60). DSRL could be of value in training programmes where supervision is difficult to provide and feedback from faculty is unavailable. In the present thesis, the potential for DSRL in off-site training in laparoscopy is discussed in Paper 1.

4. METHODOLOGICAL CONSIDERATIONS

High-quality research is based on the use of appropriate methods. The research question, aim of the study and hypotheses should determine the choice of methodology. In the following section we will describe the setting and methods used in this thesis.

Setting

The CAMES Training Programme in Laparoscopy

All of the studies included in this PhD dissertation were conducted at the Copenhagen Academy for Medical Education and Simulation (CAMES) (62). At CAMES, doctors in their first year of specialty training can participate in a basic simulation-based laparoscopic training programme. The training programme is a cross-specialty training programme for doctors working in the gynaecology, urology and surgery departments (63) and aims to prepare participants for their first supervised laparoscopic surgical procedure. The training programme is structured across two formalised training days separated by a period of self-directed training. Training takes place using VRSs and BTs. The first part of the training programme is an introductory course. This is followed by a period of self-directed training in which participants practice on VRSs and BTs. Participants train in the simulation centre, assisted by a simulator technician who can provide some feedback during training. There are two mandatory requirements in the training programme. The first requirement is to pass the TABLT (64) test on the BT and the second is to reach a predefined level of proficiency on the VRS. Gynaecologists are also required to take a theoretical test (65). When participants have completed the requirements of the training programme they are able to enrol in the final operative course. The operative course marks the end of the training programme.

Training and Assessment of Basic Laparoscopic Techniques (TABLT)

To address the need for a training and assessment tool for basic laparoscopic skills training, we developed TABLT. The tasks included in TABLT were developed for a cross-specialty training curriculum. The TABLT consists of laparoscopic tasks that test various domains of laparoscopic skills. The test includes the testing and training of ambidexterity, hand-eye coordination, accommodating for the fulcrum effect, guiding instruments via a screen, and economy of movement. Five tasks were developed covering appropriate handling of laparoscopic instruments, cutting, blunt dissection and sharp dissection. There are specific errors for each task, which are used when calculating the TABLT test score.

The tasks:

Task 1 is a coordination task. The goal of the task is to move four beads on a pegboard from one line of pegs on the right side, to another line of pegs on the left side, and back again. An error is counted if a bead is dropped. If the bead rolls outside the range of movements it is counted as two errors.

Task 2 is a cutting task, the goal of which is to cut out a circle. A circle is drawn on a soft sponge cloth and the task has been completed when the entire circle has been cut out. The circle is two millimetres wide. An error is counted when cutting outside the two-millimetre line.

Task 3 is a sharp dissection task, the goal of which is to dissect a vessel using sharp dissection technique. The vessel is made from a balloon that has two lines drawn on it. The lines are two millimetre wide and two centimetres apart. The balloon is wrapped in a soft sponge cloth. A cut into the vessel is counted as an error.

Task 4 is a blunt dissection task and its goal is to dissect a vessel using blunt dissection technique. As in Task 3, the vessel is made from a balloon that has two lines drawn on it. The lines are two millimetre wide and two centimetres apart. The balloon is wrapped in cotton wool. An error is counted if a piece of the cotton wool is ripped off, and completely removed from the rest of the task.

Task 5 is a cyst removal task, the goal of which is to remove a cyst. The cyst can also be made to simulate a gallbladder depending on the specialty. The cyst is made from two round balloons, one inside the other. The innermost balloon is filled with 60 millilitres of ultrasound gel.

The TABLT test scoring system

The scoring system is based on time and the number of errors, similar to the system used in the FLS system (48). The final score is between 0 and 708, where a higher score indicates a better result. The score is calculated by first determining a score for each task. Each task score is calculated by subtracting the time spent on a task in seconds from a maximum time of 600 seconds. The result is then divided by the average score of a group of experienced surgeons. Finally, all task scores are summed into a performance score. The score can be simply calculated using an Excel spreadsheet; the users only need to enter the number of errors and time spent completing each task.

A Mixed-Methods Approach

The papers included in the thesis are based on several research questions. Depending on the research question, an appropriate choice of methodology has been chosen, as is recommended (66).

Throughout the thesis, qualitative and quantitative methods are seen as equal and complementary (67). The first paper is a scoping review based on a systematic literature search and a qualitative thematic analysis. The second paper is a validation study. For this study, we used the unitary framework of validity, which relies on both qualitative and quantitative methods. The third was a methodological paper that explored the consequences of different standard setting methods. In this study we used a descriptive and a comparative quantitative analysis. The fourth paper was designed as a randomised controlled trial and we used quantitative statistical methods, described below. The fifth study is a mixed-methods study in which quantitative descriptive statistics were used and triangulated with findings from a qualitative content analysis of focus group interviews and individual interviews. The mixed methodology was used to explore training patterns and training methods as well as how trainees used BTs when training at home.

Validity

Validity is the process of ensuring that what you intend to measure is what is actually measured (68). With a directly observable phenomenon, such as blood glucose concentration or haemoglobin levels, validation is done within the realm of natural sciences. However, some traits such as surgical technical competency cannot be observed or measured directly. As the trait is not directly measurable, a test or an assessment tool is necessary. To ensure validity for this type of measurement, a different approach to validation is used. Frameworks have been developed for exploring the validity of tests measuring competency. These frameworks include the unitary framework of validity, although other frameworks of validity theory do also exist (69). An early validity framework used types of validity and describes the use of content, construct and criterion validity. However, the unitary framework considers all validity as a process rather than dividing it into types of validity. Hereby, validity of construct is explored; therefore, validity is construct validity. The unitary framework of validity is recommended in the ‘Standards for Educational and Psychological Testing’ (70). A newer framework called use/argument based validity also exists (71).

In this thesis we use the contemporary framework of unitary validity, an approach that has been recommended in the assessment literature for more than 15 years (72, 73). In the unitary framework of validity, validity is described as a hypothesis or process and evidence is gathered from different sources to support or negate the hypothesis (74). The sources include content, response process, internal structure, relationship to other variables and consequences of testing. When using this framework we use the following methodological terminology. A test is made of content that is a representation of the underlying construct. The construct or trait is what the test

intends to measure; for example, psychomotor skills in laparoscopic surgery. It is not the test itself that is said to be valid, but the interpretation of test scores. There is evidence to support validity when the interpretation of the test score corresponds to the construct being measured (49, 68). Evidence from *validity based on content* focuses on the content of the test and whether the content relates to the construct that the test is intended to assess (49, 68). Evidence from *validity based on response process* ensure that the intended response is elicited when administering the test. Furthermore, validity evidence from the response process includes data entry and maintaining data integrity as a means of eliminating bias that affects test scores (49, 68). Evidence from *validity based on internal structure* relates to evidence gathered from statistical analysis of the test scores to ensure reproducibility. Reliability of test scores is considered a source of validity evidence from the internal structure (49, 68). Evidence from *validity based on relationship to other variables* relates to how test score correlates to other measurements of the same construct (49, 68). Evidence from *validity based on consequences of the test* relates to the consequences of testing. This is an important source of validity evidence since potentially harmful consequences should be identified (49). Consequences as a source of validity are broad and standard setting is an important part of this. The first step in analysing the consequences of testing is to set a pass/fail level. Thereafter, the consequences of the test and the pass/fail level can be explored (75). A standard setting method is used to establish a pass/fail level.

Standard Setting

Standard setting describes the methodology used when setting pass/fail levels on a test. Throughout the development of educational theory, many different types of standard setting methods have been used (76). Nonetheless, when setting standards, the decision on a score remains a policy decision (77). Standard setting methods can be either norm-based or criterion-based, also referred to as relative and absolute (28, 76, 78). Norm-based methods are used when a pass/fail level is set according to the percentage of the students that will pass or fail. Criterion-based methods use a set criteria of passing; that is, having performed to a certain level. A criterion can be set on both surgical assessment or according to the number of correct answers that is sufficient to pass a multiple choice test. Criterion-based tests are often used to assess competency (79). Criterion-based methods are traditionally divided into examinee-centred and test-centred methods (80). Examinee-centred methods look at the examinee, determine the ability of the students and use these observations to set a pass/fail level. Test-centred methods, on the other hand, look at test characteristics, such as difficulty and relevance, and set a pass/fail level according to these. Two common methods of standard setting are the contrasting groups method (an examinee-centred

method) and the Angoff method (a test-centred method). A new method of criterion-based standard setting is the expert performance level, where the performance of experienced participants is used to set pass/fail levels. The pass/fail level is set at the mean performance level of a group of experienced participants. Although this methodology has been poorly described in educational research literature, it is used in simulation-based training (81).

Three standard-setting methods were used in this thesis. In Paper 2, the contrasting groups method is used to set a pass/fail level for the TABLT test. In Paper 3, three different methods were used to explore the consequences of the choice standard setting methodology. The three methods were the expert performance level, contrasting groups method and the Angoff method. In the expert performance level method, the pass/fail level is set at the mean performance level of a group of experienced laparoscopic surgeons. The Angoff method is a criterion-based test-centred method that, in its original form (82), consisted of asking judges to reach an agreement on the definition of a borderline student. The judges agree to define a borderline student as a student with a 50 per cent chance of passing the test. After reaching agreement on this definition, the judges would determine the performance level of a borderline student. The performance level would be described in passing percentages on each item in a test. The items scores were then averaged across different judges and a pass/fail level was set (79, 80). The Angoff method has since been modified in different ways, and sometimes includes performance data. These data are presented to the judges, who then determine passing levels through several iterations (28). There is considerable empirical evidence to support the use of Angoff method (28, 83). However, this method can provide a high pass/fail level as experts often expect too much of their students (77, 80). The Contrasting Groups method, on the other hand, is a criterion-based examinee-centred method (80). The pass/fail level is set based on a division of participants as either competent or not competent. Test scores from the participants are used and a pass/fail level is set at the intersection between the distribution of the two groups (84). The pass/fail level can be moved according to the purpose of the test, either to ensure that no competent student is failed or that no incompetent student will pass (79).

Focus Groups and Individual Interviews

Interviewing participants is a well-established method of investigation in qualitative research. Interviews can help shed light on complex problems (85). Interviews have been shown to be a particularly good methodology when exploring behaviours and experience (86). Using interviews can provide researchers with answers to 'how' and 'why' questions (87). Interviews can be analysed in different ways depending on the theoretical foundation. One method of analysis is content analysis (88), which we used in this thesis.

Focus groups are increasingly being used in healthcare research (89, 90). A focus group interview can help researchers gain insight into norms of behaviour through group interaction (90). In focus groups, relations between participants can be used to help obtain information from individuals who would not otherwise participate in an interview (91). Individual interviews, on the other hand, can help create a deeper understanding of each individual experience. Semi-structured individual interviews are often conducted based on a guide with open-ended questions (86). Individual interviews are sometimes the only source of data in qualitative research projects (85). In the present thesis, we used both focus group interviews and individual interviews. Both methodologies were used in the fifth study to help create an understanding of the use of BTs in take-home training.

Statistical Considerations

A variety of statistical methods have been used in the studies included in this thesis. Various descriptive statistics were used, as appropriate. Descriptive statistics were also presented in tables and illustrated using graphs. In the second study, reliability was measured by calculating the intraclass correlation coefficient (ICC). We used the ICC definition of single measures and absolute agreement. Pearson's correlation was used to analyse correlation. We calculated the Pearson's product moment correlation coefficient and a Pearson's r value of >0.7 was considered an acceptable degree of correlation (92). Analysis of variances (ANOVA) was used to analyse differences between groups, and a Bonferroni correction was used to allow for comparisons of more than two groups. Independent samples t-test was used to compare difference of means between groups in Papers 3 and 4. A p-value of less than 0.05 was considered statistically significant. Statistical software (SPSS, vs. 20.0 Chicago IL, USA) was used for analysis.

Ethical Considerations

All of the studies using data from participants were reported to the regional ethics committee. In accordance with Danish legislation, the regional ethics committee deemed that no approval was necessary for these studies. The studies were also reported to the Danish data protection agency and approval was given prior to commencing. The randomised controlled trial was registered at www.clinicaltrials.gov to ensure transparency. All participants were informed in writing and verbally, written consent was given before participating, and participants were informed of their rights, including their rights to withdraw their consent at any point during the studies. For the

interviews, participants were anonymised on transcription of the interviews. Ebbe Thinggaard was the only researcher who had access to non-anonymised recordings.

Overall Purpose of the Thesis

The overall purposes of thesis were to review the current knowledge of training off-site, to develop and explore validity for a training and assessment system for off-site training, to investigate the effect of take-home training in simulation-based laparoscopy programmes, and to explore the use of take-home training.

Aims

From the overall purpose of this thesis, the following aims for the studies were developed.

Study 1: Create an overview of the current knowledge of the use of BTs and the instructional designs used in off-site training programmes.

Study 2: Explore evidence of validity for a test of basic laparoscopy skills for training at home and establish a reasonable pass/fail standard.

Study 3: Determine the consequences of different standard setting methods and explore what level of competency is perceived to be adequate to begin performing supervised surgery.

Study 4: Investigate the added effects of training at home in a simulation-based laparoscopic training programme and explore the reliability of self-rating when training unsupervised.

Study 5: Describe how surgical trainees use mobile boxtrainers when training at home and explore the use of self-rating in unsupervised training.

Hypotheses

The hypotheses for the papers included in the thesis were as follows.

Study 1: Simulation-based laparoscopic training programmes for off-site training on BTs use instructional designs based on educational theory.

Study 2: Evidence of validity for a test in basic laparoscopy skills for training at home including a credible pass-fail score can be established.

Study 3: Different standard setting methods affect the consequences of a test in basic laparoscopic skills and there is a difference in the level of competency, which is perceived to be adequate to begin performing supervised surgery.

Study 4: Training laparoscopy at home has a positive effect on training and participants are able to rate their own performance using a structured self-rating system.

Study 5: Surgical trainees use mobile boxtrainers when training at home and are able to use self-rating to guide unsupervised training.

5. PRESENTATION OF THE INCLUDED PAPERS

Paper 1: Thinggaard E, Kleif J, Bjerrum F, Strandbygaard J, Gögenur I, Ritter E M, Konge L. Off-site training of laparoscopic skills, a scoping review using a thematic analysis. Surgical Endoscopy 2016;11:1-9.

Background

Simulation training is becoming a valuable addition to laparoscopic skills training in the clinical setting. Laparoscopic simulation-based training is being implemented using VRs and simple BTs. However, barriers to simulation training still pose challenges to implementation. These barriers include duty-hour restrictions, the high price of simulation equipment and the opening hours at skills labs and simulation centres. Mobile BTs may help overcome barriers to simulation training and provide trainees with the opportunity to train when they have the time. Although portable BTs may help remove the barriers associated with simulation-based laparoscopy, our knowledge of this area remains limited.

Objective

The objective of the study was to create an overview of the current knowledge on off-site training in laparoscopy.

Methods

Based on the research question and objective of the study, a search string was created. The search string was adapted and used in various online databases including: MEDLINE, ERIC, PsychINFO and Scopus. A snowballing search was also conducted and relevant websites and references from reviews on laparoscopic skills training were used to identify records. A consensus was reached regarding which records to include in the review. Two independent researchers working in collaboration did the screening. We analysed the records iteratively using a thematic analysis approach.

Results

We identified and included 22 records. Based on a thematic analysis, the following underlying themes were identified: access to training, protected training time, distribution of training, goal-setting and testing, test design, and unsupervised training. The underlying educational theories we identified included: proficiency-based learning, deliberate practice, and self-regulated learning.

Conclusions

A variety of instructional designs are used in laparoscopic skills training programmes for training on simple BTs. Instructional designs are based on different educational theories including proficiency-based learning, deliberate practice and self-regulated learning. Directed self-regulated learning could prove valuable when designing laparoscopic off-site training.

Strengths and limitations

The search identified 1978 studies, from which 22 records were included. The records varied in terms of where and how they were published. The records were published either as conference papers or articles in peer-reviewed journals. The quality and methodology of studies also varied. However, following the methodology described for scoping reviews, we did not assess the quality of the studies as each study was considered equally important (93). All records were assumed to provide findings that would help generate an understanding on the use of off-site training. The inclusion of only 22 records from a search identifying 1978 records demonstrates the broad scope of the search. We aimed to identify any source of literature that could be relevant to simulation training at home on BTs. The search was systematic in nature. A strength of our study was that we used a well-established reporting format: the STructured appROach of the Reporting In health care evaluation of Evidence Synthesis (STORIES) (94). For the analysis we used a thematic analysis approach (95).

The analytical approach shows the difference between a scoping review methodology and traditional systematic reviews used for randomised controlled trials (96). Although our study included a wide array of records and sources, our search could have been even broader. We could have included websites from relevant manufacturers of equipment for laparoscopic skills training, as well as abstract books available online from relevant conferences, websites from educational institutes and other stakeholders in surgical laparoscopic training. When deciding how broadly to search, there is always the limit of feasibility, as a broader search would have taken more time and may not have provided a better result. The records we included from websites and conference abstract books did not impact our analysis to a very high degree as these sources provide very brief descriptions of studies.

We used a thematic analysis for analysing the records (95). A benefit of thematic analysis is that it uses an open approach to existing literature and thereby allows themes to emerge from the literature. However, limitations of this approach include researcher bias and its effect on findings. Our understanding of educational theory could have influenced our findings. All of the members of

our research group were medical doctors working in surgical specialties. However, all but one of the researchers had conducted research within educational sciences and we were all familiar with a wide variety of educational theory, especially theory used in surgical training. Being familiar with the field of study helped ensure that the thematic analysis could be carried out appropriately. Familiarity with the reviewing process and educational theory created trustworthiness of our findings (97, 98).

From our analysis, we identified proficiency-based learning, deliberate practice and self-regulated learning as underlying educational theories guiding laparoscopic training programmes for off-site training. Based on our findings, we recommended that training at home in laparoscopy should be guided by sound educational principles and that the use of directed self-regulated learning should be explored further.

Paper 2: Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). British Journal of Surgery 2015;102:1106-1113.

Background

Acquiring laparoscopic techniques by training on simple BTs has shown to be an effective instructional method. Training on BTs has shown to affect both patient outcomes and improve performance on assessments. However, current training programmes have been developed for specific specialties and include advanced training such as suturing. For training to be relevant for novice laparoscopic surgeons, a new training and assessment tool is needed. The TABLT is a training tool developed for trainees to acquire the necessary laparoscopic techniques used during supervised surgery in clinical practice.

Objectives

The objectives of this study were to develop and explore evidence of validity for the TABLT test and to set a credible pass/fail level.

Methods

The TABLT test was developed in separate training programmes for basic laparoscopic skills training in surgery and gynaecology. From these training programmes, the need for a cross-specialty test emerged. The TABLT test was developed with a specific focus on construct alignment of tasks to facilitate transfer of skills into the clinical setting. To explore the validity of the TABLT test, we used the contemporary framework of validity known as the unitary framework of validity. The unitary framework of validity relies on five sources of validity evidence: content, response process, internal structure, relations to other variables, and consequences. We also established a credible pass/fail level using the contrasting groups' method.

Results

We included sixty participants in the study. Participants were doctors from surgery, gynaecology, and urology departments. Novice, intermediate and experienced laparoscopic surgeons were recruited. Novices were defined as doctors who had never performed a laparoscopic procedure. Intermediates were defined as doctors who had performed between one and 100 procedures and experienced laparoscopic surgeons were defined as doctors who had done more than 100 procedures. From the content source of validity we found that the test content was appropriate for

cross-specialty training and included a sufficient range of relevant tasks for laparoscopic skills training. From the process of scoring we found that scoring was done easily using an Excel spreadsheet and that scoring was transparent and data integrity maintained. Test scores correlated well with procedural experience, Pearson's r value was 0.73 ($p < 0.001$). We found a significant difference between groups of different levels of experience ($p < 0.001$), and scores were reliable; ICC (0.99, $p < 0.001$). These measures provided evidence validity from the internal structure of the test. As part of exploring evidence of consequences, a defensible pass/fail level was set at 358 points. At this pass/fail level, 10 per cent of novices passed the test and 10 per cent of experienced laparoscopic surgeons failed the test.

Conclusions

We found evidence of validity from five sources: content, response process, internal structure, relations to other variables and consequences. We established a credible pass/fail level for surgical trainees to reach prior to performing supervised surgery in the OR.

Strengths and limitations

A major strength of our study was that we used the unitary framework of validity (99). This is a contemporary framework recommended in the literature to establish evidence of validity for testing in surgical training (73). Previous frameworks relied on types of validity, whether it was content, construct and criterion validity, or even face validity. Face validity contributes very little to the understanding of the validity of a test (100). Types of validity only offer a limited understanding of validity and provide insufficient sources of evidence for validity in testing. In our study, we could also have chosen to use a newer framework of validity, the framework of the use-argument validity proposed by Kane (77). The use argument framework of validity has been proposed for assessment in medical education (69). The choice of methodology was based on wishing to use both a contemporary approach and one that has been used in the literature to establish validity evidence for assessments in surgical training.

We included doctors from three specialties and with different levels of experience. We included 60 participants in total. The sample size could be considered small when considering the doctors were recruited from three departments and divided into three groups. However the TABLT training was developed for cross-specialty training, which made it necessary to include doctors from different specialties. The inclusion criteria were the same for all specialties and the trait we wished to measure (basic laparoscopic technical skills) was similar across the different specialties. Using stringent inclusion criteria helped reduce the source of bias from having a wide range of doctors

participating in the study from different specialties. Having three groups of doctors also demonstrated the difference of performance in different groups. Our study included a very wide range of performance and experience levels, particularly in the group of experienced surgeons. Nonetheless, across three groups we were still able to establish a firm correlation between clinical experience, measured as the number of procedures, and performance scores on the TABLT test.

When establishing validity evidence for testing, there are two sources of bias that should be considered: the risk of construct under-representation and the risk of construct irrelevant variance (101). Looking at evidence of validity from content helps explore the bias associated with construct underrepresentation. In our study we found that the content of the TABLT test was appropriate and represented the domain of laparoscopic skills. Nonetheless, our inclusion of tasks could have been even broader. Also, the tasks could have been more advanced and included content such as suturing or camera navigation; however, suturing is not considered a simple task and would not be relevant to include in a basic laparoscopic skills test. It would also have been unreasonable to require novice laparoscopic surgeons to be able to suture prior to supervised surgery in clinical practice. Camera navigation is a relatively simple task that can easily be learned when assisting in surgery.

Bias from construct irrelevant variance can be explored both from the process and the internal structure. From the process, as a source of evidence of validity, we looked at the process of rating. The TABLT rating system was developed as a simple rating system that is easy to comprehend and record. Rating can be done using a simple spreadsheet, which allows for transparency in record keeping. Transparent recording helps maintain data integrity. To minimise the impact of different settings, all tests were done in only two places, by the same researcher. The two settings were the simulation centre and a hospital at which participants were working. Using a setting that was familiar to participants helped minimise the risk of the setting affecting the performance of the surgeons. The tests were rated on-site by the researcher and afterwards by a blinded researcher using video recordings. The blinded rater was one of the researchers from the research group, which could have affected ratings. A surgeon outside the research group could have done the ratings, but the rater would have required further training. Rater training for the TABLT test was not explored in the study.

Participants were only asked to make two attempts at the test. The first one to allow participants to familiarise themselves with the system and the second attempt was used for rating and was video-recorded. Having participants do more attempts could have created a learning curve and informed us about the learning potential of TABLT test. However, the focus of this study was on the validity of the testing aspect of the TABLT. Two attempts were considered sufficient to provide us with the data we needed in order to establish validity for the TABLT test. As a part of

establishing validity from consequences, a pass/fail level was set. The pass/fail can provide an understanding of the immediate consequences. The immediate consequences of the pass/fail level are the fail rates for experienced surgeons and passing rates for novices. Validity from consequences was analysed in this way. However, pass and fail rates only explain some of the consequences of testing.

Paper 3: Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Ensuring competency of novice laparoscopic surgeons – Exploring standard setting methods and their consequences. Journal of Surgical Education 2016.

Background

Proficiency-based simulation training in laparoscopy is growing and testing is an important part of this approach to training. In proficiency-based training, testing is used to set the minimum requirement for competency. The pass/fail level determines the minimum requirement; however, the method used to set the pass/fail has not been explored in great detail. The effect of choice of methodology on pass/fail levels should be investigated more thoroughly, as should the perceived adequacy of the pass/fail levels.

Objectives

The objectives of this study were to explore the effect of the standard setting method on the pass/fail level and to investigate whether there was a difference in the level of competency, experienced and novice laparoscopic surgeons expected a novice to reach on the TABLT test during training.

Methods

Participants were included in a validation study of the TABLT test. Participants were novice and experienced laparoscopic surgeons from surgery, gynaecology and urology departments. Each participant was asked to make two attempts on the TABLT test; the second attempt was recorded and rated on-site. The second attempt was also video-recorded and rated by a blinded rater. Three standard setting methods were used to set a pass/fail level: the expert performance level, the contrasting groups method and the Angoff method. After testing, participants were asked how high a score a novice should reach during training, how many errors were acceptable and how much time they should be allowed to spend on each task.

Results

The study included 40 participants, 20 experienced laparoscopic surgeons and 20 novices. The three standard setting methods resulted in three different pass/fail levels. The expert performance levels set the pass/fail level at 474 points, the contrasting groups method at 358 points and the Angoff method amongst experienced at 311 points and amongst novices at 386 points. The consequences of the different pass fail levels varied. The fail rate for experienced surgeons was between 0 and 50

per cent and the pass rate for novices was between 0 and 25 per cent. There was a significant difference in the level of proficiency deemed adequate by experienced and novices ($p < 0.008$). Novice laparoscopic surgeons expected novices to reach a higher score on a test during training than experienced laparoscopic surgeons did.

Conclusions

The pass/fail level of a basic test in laparoscopic skills varies highly depending on which standard-setting method is used. Experienced and novice laparoscopic surgeons have different expectations. Novices expect that other novices will be able to reach a higher test score during training than experts do.

Strengths and limitations

The focus of this study was on the choice of standard-setting method and its consequences. We examined this by using three different methods to set a pass/fail level. We calculated the fail rates amongst experienced and pass rates among novice laparoscopic surgeons to explore the consequences of the three different standard-setting methods. The aim of the study was selected on the basis of the fact that few studies have been done within surgical training on the effects of standard setting on the pass/fail level, and even fewer on the consequences. Standard setting has been explored outside the field of surgical skills training. In the literature of medical education, studies describe that the pass/level depends on the choice of standard setting method (28). These findings are in accordance with findings from our study, with the fail rate amongst experienced surgeons varying from 0 to 50 per cent. Our result illustrates a high variance in the fail rate, which depended on the choice of standard-setting method.

Choosing the right pass/fail level depends on the purpose of the test. Tests used when there are a limited number of spaces, such as an entry-test for a surgical training programme, have a different purpose than tests used for assessment during training. Testing used for formative or summative feedback should use an examinee-centred standard-setting method (28). To explore what the pass/fail should be, one must consider the level at which performance is considered sufficient. In proficiency-based training, the focus is on a high pass/fail level as trainees are expected to be not only competent but also proficient. Setting new pass/fail levels for the TABLT would require the test to correspond to a different level of training in clinical practice. A new proficiency level could be used when trainees begin to do more advanced surgery, such as when they move from partial procedures to full procedures. In this way, different pass/fail levels on a test could be explored so that the progression in clinical training corresponds to the progression in simulation training. This

would ensure a continuous use of simulation training and help trainees prepare for each step in their continued progression in the clinical setting.

Paper 4: Thinggaard E, Bjerrum F, Strandbygaard J, Konge L, Gögenur I. Take-home training facilitates distributed training, a randomized trial. Surgery 2016; Submitted.

Background

Acquiring surgical skills during simulation training is becoming a valuable addition to clinical training in the OR. Laparoscopic techniques can be acquired in simulation-based training using simple box trainers, which enable training at home. Take-home training helps trainees overcome barriers and makes training accessible. Although the feasibility of take-home training has been investigated, the effect of training at home has not been explored sufficiently. It is necessary to explore the effects of take-home training. Training at home also presents the challenge of missing feedback when training unsupervised. Rating their own performance may provide participants with a type of feedback, but the reliability of self-rating in laparoscopic training has not been thoroughly investigated.

Objectives

The objectives of this study were to investigate the added benefit of training at home in a simulation-based training programme and to explore if trainees were able to rate their own performance on the TABLT test.

Methods

We designed and conducted a randomised controlled trial with a blinded rater. Participants were doctors in their first year of training. Participants were recruited during a basic laparoscopic skills programme in which VRSs and BTs are used. After inclusion, participants were randomised to either having access to take-home training or following the regular course with only on-site training. Participants used logbooks to record their training. Training patterns were measured, including the time to complete the training course, the time spent on training and the number of training sessions.

Results

Thirty-six participants were included in the study; 18 in each arm. We found a significant difference in the number of training sessions (5.8 versus 2.3, $p < 0.001$). We found no difference in the time to complete the training programme (86 vs. 89 days, $p = 0.89$), the time spent on training (302 vs. 218 minutes, $p = 0.26$) or the performance score (493 vs. 460, $p = 0.07$). Participants were able to rate their own performance using video recordings. Self-rating provided reliable ratings correlating well

to those of a blinded rater; ICC 0.86, $p < 0.001$.

Conclusions

When free to choose, participants choose to train in a distributed manner. They divide their training into shorter sessions and they train more frequently. They do not take longer to complete a training course and do not spend more time on training. Participants in a laparoscopic skills course are able to rate their own performance using a simple scoring system. Ratings from self-ratings are reliable.

Strengths and limitations

This study aimed to investigate the added benefit of training at home in a simulation-based laparoscopic training course. When designing the study we measured three equivalent outcomes: time to complete the training course, the time spent training, and the number of training sessions. For the sample size calculation we chose to use time to complete the training course. We anticipated that training at home would affect training patterns and make it possible for trainees to pass a test earlier. Training at home would reduce the length of the training course and allow trainees to start performing supervised surgery earlier. We also anticipated that participants would reach a higher proficiency level as they had more access to training. However, the length of the training course was determined by the mandatory elements of the course. The course started with an introduction course and ended with the operative course (63). Therefore, the length of the training remained unaffected. We also found no significant statistical difference in the level of proficiency. Although there was a tendency for the scores of participants training at home to be higher, we were unable to conclude that this was due to the access provided in take-home training. Our finding corresponds well with the fact that participants were able to rate their own performance. Self-rating helped participants reach the required pass/fail level, but also acted as a minimum requirement for trainees to reach.

One limitation of our study design is the sample size calculation. For a randomised controlled clinical trial, only one measurement can be chosen for the sample size calculation. At the end of the study inclusion period of one year, we did a new sample size calculation to evaluate the feasibility of continuing the trial. The new sample size calculation showed us that more than 11,000 participants would be required to measure the difference using time to complete the course. Based on this calculation it was not feasible to continue the trial. Furthermore, the trial had already yielded significant results about what had been chosen as secondary outcomes. These significant results in combination provided information on the distribution of training patterns amongst trainees. We designed the study as a randomised controlled trial set in an actual simulation-based laparoscopy

course. As we conducted the trial in a real setting, we were unable to control all variables. Thus, we pushed the boundaries of what is possible to investigate using a randomised controlled trial design. Despite the challenges with variables such as the on-going clinical training and the use of both VRSs and BTs, we still found a significant difference when comparing the two groups. Using a real laparoscopic course setting has helped us gain information on the distribution and training patterns chosen by participants, and has also provided valuable information about how home training affects participant behaviour in a real laparoscopic skills course. Trainees divided their training into shorter and more frequent intervals. Participants opted for distributed training when training at home. However, these results do not address how and why the participants trained. To further explore these questions, a study with a different choice of methodology would be needed. ‘How’, ‘why’ and ‘what’ questions are in focus in medical education (94), but these questions were outside the scope of our study.

**Paper 5: Thinggaard E, Konge L, Bjerrum F, Strandbygaard J, Gögenur I, Spanager L.
Take-home training in a simulation-based laparoscopy course. Surgical Endoscopy 2016.**

Background

It has been established that simulation-based training is effective for acquiring laparoscopic skills. Nevertheless, implementation of simulation-based training, particularly implementation of training that is accessible to trainees, has been slow to spread. Training at home has been shown to be feasible, but there has been little research into how and why trainees train as they do when training at home. Even fewer studies have addressed the need of feedback in unsupervised training and the role of self-rating as a mean of providing guidance in unsupervised training.

Objectives

The objectives of this mixed-methods study were to describe the used of BTs when training at home and to explore the training patterns, and how participants trained and what they trained on.

Methods

The study was designed as a mixed-methods study in which we incorporated methodologies from both quantitative research traditions and qualitative research traditions. The study consisted of two equally important parts: a descriptive quantitative analysis followed by a qualitative analysis of data from interviews. The participants were recruited amongst doctors on a basic laparoscopic skills course. All participants were offered training at home on mobile BTs and training at a simulation centre on VRSs and BTs. Participants were given logbooks to record when and how they trained. Data from logbooks were analysed using descriptive statistics and graphs. Focus groups and individual interviews were held and analysed by researchers following a content analysis methodology.

Results

We included 18 participants in the study. From the quantitative analysis we found that participants used an individualised approach to training. This finding was supported in the qualitative data, where participants described that they trained the task they wanted to train in the sequence they deemed most appropriate. Participants also mixed their training methods by training on both BTs and VRSs. They also mixed their use of training settings by training both at home and at the simulation centre. Participants integrated the possibility of training in a specific location at the simulation centre with the flexibility of training at home. Findings from the quantitative data

analysis were used to direct the qualitative analysis. The qualitative analysis was conducted on data from focus groups and individual interviews. We identified the following themes: training method, training pattern, feedback and self-regulation.

By way of conclusions, we found that training patterns varied and trainees used an individualised approach. Mandatory training strongly affected training patterns and testing provided this structure. We also found that self-rating helped guide participants during unsupervised training.

Strengths and limitations

This study was designed as a mixed-methods study. We chose to use methods from quantitative research traditions as well as qualitative research traditions. The mixed-methods methodology has been used in healthcare-related research as well as medical educational research (102). The quantitative part of our research was a descriptive statistical analysis, while the qualitative part of the study was an analysis of focus group and individual interviews. Selecting a mixed methodology enabled us to consider the weaknesses and strength of the study from both a quantitative approach and qualitative approach.

With regard to quantitative research, traditional sources of bias include internal validity, external validity, generalisability, reliability, and objectivity (97). The sample size in our study only consisted of 18 participants, which is a small sample size. The small sample size affects the external validity of the study and makes it more difficult to generalise the findings to another population. Despite the small sample size, we were able to include a high variety of participants. Participants were both female and male, and were doctors from three different specialties, working at hospitals in various distances to the simulation centre. Participants were primarily novices and had different levels of experience with laparoscopic surgery. Having a variety of participants helps generalise the findings to other settings. The study was a descriptive study and we used descriptive statistics. This type of study is explorative and cannot provide confirmation of observations. A blinded randomised controlled study would have been better suited in order to establish confirmatory findings. However, the aim of the study was to explore the use of take-home training and the methodology, so the sample was appropriately aligned with these aims. Bias related to objectivity included considering the role of the researcher in our project. The primary researcher had different roles, both as a faculty member present at the final test and also as the facilitator of the interviews. Therefore, the researchers' role was a source of bias that may have affected our findings. Within the qualitative research traditions, however, bias is treated in terms of trustworthiness (97, 98). To help provide transferability, it is recommended that researchers have an understanding of the context and

that there is a sufficient length and number of data collection and collection sessions (98). The qualifications of the research group were appropriate. The research group behind the study consisted of researchers who were familiar with the use of qualitative research methodology, as well as researchers who were well versed in research in surgical skills training. This helped improve and ensure the transferability of knowledge generated from this study. Another source of evidence of trustworthiness is transferability, which includes reflective practice and having a transparent audit trail (98). Data collection was documented thoroughly to ensure an appropriate audit trail. To help improve the quality of the study we used different qualitative methodologies, both focus group interviews and individual interviews. Different sources of data helped ensure sufficient data for data triangulation and saturation, which is an important source of trustworthiness (97, 98). Iterative questioning was used during focus group and individual interviews, which improved the confirmability and credibility. An iterative questioning also ensured that meaning was explicit. Based on these control mechanisms, which are part of the methodological framework in quantitative research, there is ample evidence of trustworthiness to support the findings from our qualitative data analysis.

6. DISCUSSION

The overall purposes of the thesis were to review the current knowledge on training off-site in laparoscopy, develop and explore validity for a training and assessment system for off-site training, to investigate the effect of take-home training in a simulation-based laparoscopic training programme, and to explore the use of take-home training. We have found (1) that training programmes for off-site training are based on a variety of instructional designs and educational theory, (2) that the TABLT test is supported by evidence of validity, (3) that the choice of standard setting method affects the pass/fail level of a test, (4) that trainees use an individualised and distributed approach to training when training at home, (5) that training requirements and testing are determinants of training patterns, and (6) that self-rating is reliable and helps provide guidance when training unsupervised. These findings are discussed below.

Educational Theory and Take-Home Training

In the first paper in this thesis we conducted a review of the literature on off-site training and found that a variety of educational theories and instruction designs are used in off-site laparoscopic training programmes. However, the educational theories or conceptual frameworks were rarely stated. It has been recommended that the conceptual framework is stated clearly and described when conducting research in medical education, although this is often lacking in published literature (103). For research in medical education, the theoretical framework is an essential part of establishing evidence in support of findings (44). Furthermore, it is recommended that educational theory is used as a foundation for successful curricular designs (26, 104). In our review we found that educational theory used in training programmes for off-site training included deliberate practice, proficiency-based learning and self-regulated learning. Extant literature has recommended deliberate practice for acquisition of technical skills. Deliberate practice relies on the use of well-defined tasks, distributed learning and immediate feedback (57). In off-site training it is a challenge to implement immediate feedback, as faculty is not readily available. In Paper 4 and Paper 5 in this thesis we asked trainees to use self-rating as a method of getting feedback. In Paper 4, self-rating was shown to be reliable and in Paper 5 we found that self-rating was seen as a useful source of feedback that helped guide unsupervised training. We also found that trainees used a distributed training pattern and that there was a significant difference in the training pattern. Participants training at home trained more often and in shorter intervals than participants who trained only at the simulation centre. This finding was confirmed in Paper 5 and participants described how they used an individualised approach by distributing training according to their schedule. Distributed training

has shown to be effective and to improve retention (52). Another way to improve retention is through testing. In Paper 5 we found that testing and mandatory training requirements determine training patterns. Furthermore, that self-rating was seen as a help in self-regulating in unsupervised training. In the review, we concluded that principles from the DSRL could help guide curricular design for simulation-based laparoscopic training off-site. DSRL could be used to help guide an independent learning approach in unsupervised training. We did this by using self-rating to support the structural framework in Papers 4 and 5 (59, 60). We found that self-rating provided the needed structure to guide unsupervised training. DSRL is an emerging educational framework and its use remains to be shown in laparoscopic skills training. Recently, DSRL has been used in combination with principles of mastery learning and has been shown to be cost-effective (105). Mastery learning has recently received renewed attention in medical education literature and has been recommended (106-108). In laparoscopy, one study showed that mastery learning was effective when implemented in a curriculum for laparoscopic inguinal hernia repair (20). Laparoscopic training programmes have often been based on proficiency-based training that requires a higher level of performance from trainees than competency-based training (27). In mastery learning, however, participants are expected to reach consecutive levels of mastery during training. Mastery learning could be used in take-home training in laparoscopy. However, mastery learning requires high-quality assessment supported by evidence of validity in order to be effective (106). To implement principles of mastery learning it is necessary to develop a series of assessments. Also, mastery learning requires that current standard setting methods for testing are modified (109). The need to modify standard setting methods shows the importance of using educational theory for curricular designs. For a off-site training curricula to be effective it should be based on sound educational theory so that content, assessments and clinical practice are aligned appropriately.

Consequences of Testing and Validity

Testing and setting a pass/fail level can be used to help trainees prepare for clinical practice. Testing is not only important in mastery learning, but is also a part of proficiency-based training and also used in summative feedback (77). In Paper 3 we explored the immediate consequences of testing and the choice of standard setting method. We found that different standard setting methods resulted in different pass/fail levels. This finding was supported in the literature on standard setting in medical education research (28). When implementing tests it is important that they are supported by a wide range of validity evidence (49, 72). The analysis of consequences of tests and their pass/fail levels is a source of evidence for validity in the unitary framework of validity (99). In Paper 2 we found evidence in support of validity for the TABLT test including evidence from the

source of consequences of testing. Nonetheless, the analysis of the immediate consequences of the pass/fail level is only part of the exploration of validity (75). Validity from the consequences of a test can include consequences to the trainees, faculty, patients and society in general. Consequences to the trainee are particularly important for a basic laparoscopic skills test used prior to supervised surgery of patients. Positive consequences of passing include a positive learning experience and increased confidence. However, if the pass/fail level is set too low, confidence may turn into overconfidence as participants have not acquired the necessary level of competency. Failing has both positive and negative consequences. Negative consequences can include a lack of progress for trainees and trainees having to face the fact that they are not yet proficient. However, these consequences are outweighed by the need for patient safety. However, failing can also help participants to reflect on their level of competency and identify areas that need improvement. To address these consequences of testing, a summative test could be accompanied by formative feedback. The consequences of testing are an important part of establishing evidence of validity and are the focus in the use/argument framework of validity proposed by Kane (71). The use/argument approach validity framework highlights the decisions and the inferences made from test scores (69). In this framework it is important to clarify what inferences and which decisions will be made from test scores (71). Inferences from test scores could include insight into the proficiency level in laparoscopic surgery and decisions could include an evaluation of a trainee's readiness to perform unsupervised surgery in a clinical setting. The use/argument approach relies on an exploration of scoring, generalising, exploration and implications (71). The use/argument has been used recently to explore validity for the Objective Structured Assessment of Technical Skills in current literature (110). In both the unitary framework of validity and the use/argument approach, validity is seen as a continuous process. Validity is seen as a process in which evidence is gathered to support validity (70). In the second paper in this thesis we used the unitary framework of validity. However, the TABLT test has since been implemented in a cross-specialty course. After the implementation, new consequences have emerged and the use/argument approach could be used to further explore validity evidence for the TABLT test. The focus on the analysis of consequences for testing demonstrates the importance of the effect of testing on training.

Training Patterns and Distributed Learning.

When participants trained without supervision, we found that they adopted an individualised and distributed training pattern. Distributed practice is an improvement on massed practice and has been recommended for laparoscopic skills training (27, 104). Distribution of training improves retention (52, 53), although the optimal frequency of training has not been determined (56). In Paper 4 we

found that participants use a distributed training pattern. This was supported in Paper 5. The descriptive statistical analysis revealed a tri-phasic training pattern. The first phase was one of moderate training intensity, the second phase was a period of low-intensity training and the third phase was one of high-intensity training. We were able to triangulate these findings using quantitative and qualitative data analysis. Training patterns were discussed in detail during the focus group interviews and individual interviews. In the qualitative analysis we looked at why this training pattern emerged. In the first part of the tri-phasic training pattern, participants had to overcome the initial challenge of setting up the system. They overcame this challenge by being motivated by having access to training, which also helped them adopt a moderate level of training intensity. However, training intensity was reduced during the second phase of the training, which could be seen as a lack of internal motivation. Due to the requirements in clinical work placing and other time constraints, internal motivation was insufficient for maintaining a moderate training intensity. In the third phase of training, participants had a high level of training intensity. This high level of training intensity was described as being associated with the approaching test deadline. Participants were freely able to organise their testing sessions and could come for testing whenever they pleased. Nonetheless, 80 per cent of participants decided to go for testing within a week of the operative course. This finding suggests that deadlines posed by testing or other mandatory training elements provided external motivation to train. Training requirements ensure that participant train to reach a sufficient level of competency.

Testing and mandatory training requirements

Testing is said to be a driver of learning (111) and is an important part of competency-based medical education, particularly proficiency-based training. Although testing is a driver of learning, learning can vary from trainee to trainee and depends on several other factors, such as internal motivation (112). Despite this variation, testing has shown to improve learning outcomes (47). Testing in laparoscopy has been introduced as part of proficiency-based training (27) and the effects of proficiency-based training have been demonstrated (29). In Paper 1 we found that many instructional designs for off-site training included testing. Furthermore, it has been recommended to implement testing in off-site training programs (113). In Paper 5 we found that testing and mandatory training are determinants of training patterns. Participants trained according to their needs and external requirements such as mandatory testing. From the logbooks and interviews we found that participants trained in a tri-phasic training pattern, and the highest training intensity was seen when the deadline to pass a test approached. Although many other sources of motivation exist, our study showed that testing can help ensure that trainees reach a necessary level of technical

competency. Implementing testing in a off-site simulation-based training program can help trainees prepare for supervised surgery on patients. Testing will hereby ensure that both trainees and patients benefit.

Unsupervised Training and Self-Rating

Take-home training is challenged by the fact that feedback from faculty is not readily available in unsupervised training. As an alternative, we asked the participants to rate their performance. In Paper 5 we explored the role of self-rating. Feedback and self-rating was discussed in the interviews and self-rating was mentioned as a source of feedback. Self-rating was described as being useful especially when preparing for testing. We believe that self-rating is perceived in this way because the testing system is relatively simple but still relevant for novice laparoscopic surgeons. It is easy for participants to interpret the scores as they are based on the number of errors and time. While some may criticise this as a simplification of the skills needed in laparoscopic surgery, the TABLT test was developed for basic technical skills. Simple numeric feedback based can provide useful feedback when training to pass a test (114). However, narrative feedback has shown to be effective for complex skills training (115, 116) and we suggest that narrative feedback from supervising surgeons could be implemented when simulation training is used for more advanced training. Narrative feedback will also help provide participants with information on other skills that are relevant to laparoscopic surgery, such as how to ensure proper ergonomics when working. In our study, self-rating allowed trainees the freedom to choose when to get feedback and this has shown to be effective as participants are in control of when they receive feedback (117). In Paper 4 we found that participants saw self-rating as a useful method with which to guide their training and was useful for directing self-regulation. The use of DSRL for laparoscopic training programmes off-site was discussed in Paper 1 and we concluded that this could be a potentially useful educational theory for off-site training programmes. DSRL is based on SRL which is important in unsupervised training and is recommended in a recent systematic review on simulation-based training (61). DSRL is when the focus is on autonomous learning combined with the need to guide trainees in their learning experiences (60). Faculty can provide facilitation and a framework in SRL by using principles of DSRL (59). In Paper 5 we could see that the principles DSRL were in use. Self-rating was used to provide guidance in unsupervised training. Participants were able to self-regulate within a structured framework. This was described as an individualised approach that was structured by mandatory training requirements.

7. Conclusions

Laparoscopic surgery requires very specific skills. Basic skills can be acquired through simulation-based training outside the OR. Simulation-based training is an effective method of instruction, but barriers to simulation training exist and implementation remains to be explored. Take-home training has been suggested as a solution to overcome barriers and we have explored this topic in the present thesis. We have created an overview of the literature on off-site training, developed and explored validity for a test for off-site training, looked at the consequences of the pass/fail levels for this test, and – through a randomised controlled trial and mixed-methods study – found that off-site training allows for distributed training. Furthermore, we established that self-rating is reliable, that self-rating help guide unsupervised training, and that mandatory training requirements and testing are determinants of training patterns. From these findings we conclude that training off-site is effective when there is access to training, guidance during training and mandatory training requirements.

8. Implications and Future Research

The findings presented in this thesis offer several practical implications. The TABLT test has been implemented in a simulation-based cross-specialty laparoscopic training programme (63), the TABLT test has been used in research, to explore the use of laser visual guidance (118) and is being used in an on going research project to investigate the feasibility of self-certification. Furthermore, the TABLT test and findings from the thesis have some potential future implications. The TABLT test could be used in current cross-specialty laparoscopic training courses and passing the test could be implemented as a prerequisite for supervised surgery on patients. Our results regarding distributed training patterns indicate that laparoscopic training courses should move away from one-day courses and boot camps, and instead use an approach that encourages distributed practice.

Our findings and the discussion thereof have shown some of the gaps in current literature and can help guide new research in this area. Off-site training still remains to be explored in more depth, especially research that aims to explore the use in settings where simulation training is not accessible to trainees. Research should look at the role off-site training as long distance courses using internet-based simulation courses. Feedback should be a focus for this type of research especially for advanced laparoscopic skills. How best to implement feedback in off-site

training should be examined. Investigating off-site training is of particular importance in countries where laparoscopy is still being implemented.

A limitation in our studies is that we did not explore the transfer of skills from training on TABLT into the clinical setting. Future research should investigate this more thoroughly. Research should aim to explore the use of mastery learning and different standard-setting levels and how this affects performance. Different pass/fail levels could be established for corresponding levels of clinical advancement and the effect of continued simulation training throughout the clinical progression of trainees should be investigated. Simulation training has a role not only in preparing trainees for clinical practice but also as a supplement for technical skills training in clinical practice. Furthermore, laparoscopic training also has a potential in maintaining technical skills and this aspect of simulation-based training remains to be explored in more depth.

Simulation-based training in laparoscopy has become an indispensable part of how trainees acquire laparoscopic skills. However, for simulation-based training to be effective we need to explore its role continuously. As the implementation of laparoscopic training spreads we need to move outside single institutions and simulation centres and conduct multiple centre research on a national and international level. National and international collaboration will help gain further insight in how to optimize simulation-based training, how best to equip our trainees and most importantly how to ensure that our patients receive the best surgical care we can deliver.

9. REFERENCES

1. SPANER SJ, WARNOCK GL. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *Journal of Laparoendoscopic & Advanced Surgical Techniques*. 1997;7(6):369-73.
2. Keus F, de Jong JA, Gooszen HG, van Laarhoven CJ. Laparoscopic versus open cholecystectomy for patients with symptomatic cholecystolithiasis. *The Cochrane database of systematic reviews*. 2006(4):CD006231.
3. Sauerland S, Jaschinski T, Neugebauer EA. Laparoscopic versus open surgery for suspected appendicitis. *The Cochrane database of systematic reviews*. 2010(10):CD001546.
4. Darzi SA, Munz Y. The impact of minimally invasive surgical techniques. *Annual review of medicine*. 2004;55:223-37.
5. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *American journal of surgery*. 1998;175(6):482-7.
6. Dent TL. Training, credentialling, and granting of clinical privileges for laparoscopic general surgery. *American journal of surgery*. 1991;161(3):399-403.
7. Club. TSS. A Prospective Analysis of 1518 Laparoscopic Cholecystectomies. *New England Journal of Medicine*. 1991;324(16):1073-8.
8. Soper NJ, Fried GM. The fundamentals of laparoscopic surgery: its time has come. *Bulletin of the American College of Surgeons*. 2008;93(9):30-2.
9. Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Annals of surgery*. 2013;257(4):586-93.
10. Zendejas B, Cook DA, Brydges R, Hamstra SJ, Bingener J. Comparative effectiveness of virtual reality versus box trainer simulators for laparoscopic surgery training: A systematic review & meta-analysis. *Journal of surgical research*. 2013;2(179):188.
11. Partridge R, Hughes M, Brennan P, Hennessey I. There is a worldwide shortfall of simulation platforms for minimally invasive surgery. *Journal of Surgical Simulation*. 2015;2:12-7.
12. Caban AM, Guido C, Silver M, Rossidis G, Sarosi G, Ben-David K. Use of collapsible box trainer as a module for Resident Education. *J Soc Laparoendo*. 2013;17:440-4.

13. Lau W, Leow C, Li AK. History of endoscopic and laparoscopic surgery. *World journal of surgery*. 1997;21(4):444-53.
14. Tanis PJ, Buskens CJ, Bemelman WA. Laparoscopy for colorectal cancer. *Best practice & research Clinical gastroenterology*. 2014;28(1):29-39.
15. Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: Ready for widespread use? *Surgery*. 2004;135(1):28-42.
16. Gue S. Home-made videoscopic trainer for operative laparoscopic surgery. *The Australian and New Zealand journal of surgery*. 1995;65:820-1.
17. Coleman J, Nduka CC, Darzi A. Virtual reality and laparoscopic surgery. *British Journal of Surgery*. 1994;81(12):1709-11.
18. Brown W, Satava R, Rosen J. Virtual reality and surgical training: simulating the future. *Minimally Invasive Therapy*. 1994;3(2):81-6.
19. Satava RM. Virtual reality surgical simulator. The first steps. *Surgical endoscopy*. 1993;7(3):203-5.
20. Zendejas B, Cook DA, Hernandez-Irizarry R, Huebner M, Farley DR. Mastery learning simulation-based curriculum for laparoscopic TEP inguinal hernia repair. *Journal of surgical education*. 2012;69:208-14.
21. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*. 2002;236(4):458-63; discussion 63-4.
22. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, 3rd, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *American journal of surgery*. 2007;193(6):797-804.
23. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, et al. Effect of virtual reality training on laparoscopic surgery: Randomised controlled trial. *British medical journal*. 2009;338:1253.
24. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *The British journal of surgery*. 2004;91:146-50.
25. Zevin B, Aggarwal R, Grantcharov TP. Surgical simulation in 2013: why is it still not the standard in surgical training? *Journal of the American College of Surgeons*. 2014;218(2):294.

26. Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. *The New England journal of medicine*. 2006;355(25):2664-9.
27. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Annals of surgery*. 2005;241(2):364-72.
28. Cusimano MD. Standard setting in medical education. *Academic medicine : journal of the Association of American Medical Colleges*. 1996;71(10 Suppl):S112-20.
29. Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. *American journal of surgery*. 2010;199(1):115-20.
30. Zendejas B, Ruparel RK, Cook DA. Validity evidence for the Fundamentals of Laparoscopic Surgery (FLS) program as an assessment tool: a systematic review. *Surgical endoscopy*. 2015:1-9.
31. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P. Tablet based simulation provides a new solution to accessing laparoscopic skills training. *Journal of surgical education*. 2013;70:161-3.
32. Arden D, Hacker MR, Jones DB, Awtrey CS. Description and Validation of the Pelv-Sim: A Training Model Designed to Improve Gynecologic Minimally Invasive Suturing Skills. *Journal of minimally invasive gynecology*. 2008;15:707-11.
33. Chandrasekera SK, Donohue JF, Orley D, Barber NJ, Shah N, Bishai PM, et al. Basic Laparoscopic Surgical Training: Examination of a Low-Cost Alternative Netherlands: Elsevier; 2006;6:1285-91.
34. Robinson JK, Kushner DM. Development and validation of a home-based, mirrored, gynecologic laparoscopy trainer. *Journal of minimally invasive gynecology*. 2006;13(2):102-7.
35. Bruynzeel H, De Bruin AFJ, Bonjer HJ, Lange JF, Hop WCJ, Ayodeji ID, et al. Desktop simulator: Key to universal training? *Surgical endoscopy*. 2007;21:1637-40.
36. Xiao D, Jakimowicz JJ, Albayrak A, Buzink SN, Botden SM, Goossens RH. Face, content, and construct validity of a novel portable ergonomic simulator for basic laparoscopic skills. *Journal of surgical education*. 2014;71(1):65-72.
37. Chung SY, Landsittel D, Chon CH, Ng CS, Fuchs GJ. Laparoscopic skills training using a webcam trainer. *The Journal of urology*. 2005;173:180-3.

38. Moreira-Pinto J, Silva JG, de Castro JLR, Correia-Pinto J. Five really easy steps to build a homemade low-cost simulator. *Surgical innovation*. 2013;20(1):95-9.
39. Blirup-Jensen D, Gogenur I, Rosenberg J. New model for training laparoscopic skills. *Ugeskrift for laeger*. 2005;167:3780-1.
40. Pokorny MR, McLaren SL. Inexpensive home-made laparoscopic trainer and camera. *ANZ journal of surgery*. 2004;74:691-3.
41. Johnston TJ, Tang B, Alijani A, Tait I, Steele RJ, Ker J, et al. Laparoscopic surgical skills are significantly improved by the use of a portable laparoscopic simulator: results of a randomized controlled trial. *World journal of surgery*. 2013;37:957-64.
42. Ruparel RK, Brahmabhatt RD, Dove JC, Hutchinson RC, Stauffer JA, Bowers SP, et al. "iTrainers"--novel and inexpensive alternatives to traditional laparoscopic box trainers. *Urology*. 2014;83(1):116-20.
43. Alfa-Wali M, Antoniou A. Eco-friendly laparoscopic home trainer. *Simulation in Healthcare*. 2011;6:176-9.
44. Prideaux D, Bligh J. Research in medical education: asking the right questions. *Medical education*. 2002;36(12):1114-5.
45. Flexner A. Medical education in the United States and Canada. *Bulletin of the World Health Organization*. 2002;80(7):594-602.
46. Cox M, Irby DM, Cooke M, Irby DM, Sullivan W, Ludmerer KM. American medical education 100 years after the Flexner report. *New England journal of medicine*. 2006;355(13):1339-44.
47. Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Medical education*. 2009;43(1):21-7.
48. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surgical endoscopy*. 2003;17(6):964-7.
49. Downing SM. Validity: on meaningful interpretation of assessment data. *Medical education*. 2003;37(9):830-7.
50. Donovan JJ, Radosevich DJ. A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*. 1999;84(5):795.
51. Shea CH, Lai Q, Black C, Park J-H. Spacing practice sessions across days benefits the learning of motor skills. *Human movement science*. 2000;19(5):737-60.

52. Benjamin AS, Tullis J. What makes distributed practice effective? *Cognitive Psychology*. 2010;61(3):228-47.
53. Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychol Sci Public Interest*. 2013;14(1):4-58.
54. Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. *Surgical endoscopy*. 2002;16(6):957-61.
55. Spruit EN, Band GP, Hamming JF. Increasing efficiency of surgical training: effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surgical endoscopy*. 2014;29(8):2235-2243
56. De Win G, Van Bruwaene S, De Ridder D, Miserez M. The optimal frequency of endoscopic skill labs for training and skill retention on suturing: A randomized controlled trial. *Journal of surgical education*. 2013;70(3):384-93.
57. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic medicine : journal of the Association of American Medical Colleges*. 2004;79(10 Suppl):S70-81.
58. Saks K, Leijen Ä. Distinguishing Self-directed and Self-regulated Learning and Measuring them in the E-learning Context. *Procedia-Social and Behavioral Sciences*. 2014;112:190-8.
59. Brydges R, Dubrowski A, Regehr G. A new concept of unsupervised learning: directed self-guided learning in the health professions. *Academic medicine : journal of the Association of American Medical Colleges*. 2010;85(10 Suppl):S49-55.
60. Brydges R, Nair P, Ma I, Shanks D, Hatala R. Directed self-regulated learning versus instructor-regulated learning in simulation training. *Medical education*. 2012;46(7):648-56.
61. Brydges R, Manzone J, Shanks D, Hatala R, Hamstra SJ, Zendejas B, et al. Self-regulated learning in simulation-based training: A systematic review and meta-analysis. *Medical education*. 2015;49(4):368-78.
62. Konge L, Ringsted C, Bjerrum F, Tolsgaard MG, Bitsch M, Sorensen JL, et al. The Simulation Centre at Rigshospitalet, Copenhagen, Denmark. *Journal of surgical education*. 2015;72(2):362-5.
63. Bjerrum F, Sorensen JL, Thinggaard E, Strandbygaard J, Konge L. Implementation of a Cross-specialty Training Program in Basic Laparoscopy. *JSLs : Journal of the*

- Society of Laparoendoscopic Surgeons / Society of Laparoendoscopic Surgeons. 2015;19(4).
64. Thinggaard E, Bjerrum F, Strandbygaard J, Gogenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *The British journal of surgery*. 2015;102(9):1106-13.
 65. Strandbygaard J, Maagaard M, Larsen CR, Schouenborg L, Ottosen C, Ringsted C, et al. Development and validation of a theoretical test in basic laparoscopy. *Surgical endoscopy*. 2013;27(4):1353-9.
 66. Ringsted C, Hodges B, Scherpbier A. 'The research compass': an introduction to research in medical education: AMEE Guide no. 56. *Medical teacher*. 2011;33(9):695-709.
 67. Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet*. 2001;358(9280):483-8.
 68. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *The American journal of medicine*. 2006;119(2):166 e7-16.
 69. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Medical education*. 2015;49(6):560-75.
 70. American Educational Research A, American Psychological A, National Council on Measurement in E, Joint Committee on Standards for E, Psychological T. Standards for educational and psychological testing. 2014.
 71. Kane M, Brennan R. Educational measurement. Validation, 4th ed. Westport: American Council on Education/Praeger. 2006; 17-64
 72. Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. *Annals of surgery*. 2015;261(2):251-62.
 73. Korndorffer JR, Jr., Kasten SJ, Downing SM. A call for the utilization of consensus standards in the surgical education literature. *American journal of surgery*. 2010;199(1):99-104.
 74. Cook DA. When I say... validity. *Medical education*. 2014;48(10):948-9.
 75. Cook DA, Lineberry M. Consequences Validity Evidence: Evaluating the Impact of Educational Assessments. *Academic medicine*. 2016; 91;6:785-795.

76. Jaeger RM. Certification of Student Competence. In: Linn RL, editor. Educational Measurement. 3rd edn. ed. New York: American Council on Education and Macmillian; 1989. 485-514.
77. Kane MT, Crooks TJ, Cohen AS. Designing and Evaluating Standard-Setting Procedures for Licensure and Certification Tests. *Advances in health sciences education : theory and practice*. 1999;4(3):195-207.
78. McKinley DW, Norcini JJ. How to set standards on performance-based examinations: AMEE Guide No. 85. *Med Teach*. 2014;36(2):97-110.
79. Norcini JJ. Setting standards on educational tests. *Medical education*. 2003;37(5):464-9.
80. Ben-David MF. AMEE Guide No. 18: Standard setting in student assessment. *Medical teacher*. 2000;22(2):120-30.
81. Kolkman W, Van De Put MAJ, Wolterbeek R, Trimbos JBMZ, Jansen FW. Laparoscopic skills simulator: Construct validity and establishment of performance standards for residency training. *Gynecological Surgery*. 2008;5(2):109-14.
82. Angoff WH. Scales, norms and equivalent scores. In: Thorndike RL, editor. Educational Measurement. 2nd edn. ed. Washington: American Council on Education; 1971. 508-600.
83. Norcini JJ, Shea JA. The credibility and comparability of standards. *Appl Meas Educ*. 1997;10(1):39-59.
84. Clauser BE, Clyman SG. A contrasting-groups approach to standard setting for performance assessments of clinical skills. *Academic medicine : journal of the Association of American Medical Colleges*. 1994;69(10 Suppl):S42-4.
85. DiCicco-Bloom B, Crabtree BF. The qualitative research interview. *Medical education*. 2006;40(4):314-21.
86. Britten N. Qualitative research: qualitative interviews in medical research. *British medical journal*. 1995;311(6999):251-3.
87. Kuper A, Reeves S, Levinson W. An introduction to reading and appraising qualitative research. *Bmj*. 2008;337(7666):404-7.
88. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qualitative health research*. 2005;15(9):1277-88.
89. Danielsen AK, Spanager L. Anvendelse af fokusgruppeinterview inden for sundhedsvidenskab. *Ugeskr Læger*. 2012;174:19.

90. Kitzinger J. Qualitative research. Introducing focus groups. *British medical journal*. 1995;311:299.
91. Malterud K. The art and science of clinical knowledge: evidence beyond measures and numbers. *The Lancet*. 2001;358(9279):397-400.
92. Downing SM. Reliability: on the reproducibility of assessment data. *Medical education*. 2004;38(9):1006-12.
93. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice*. 2005;8(1):19-32.
94. Gordon M, Gibbs T. STORIES statement: Publication standards for healthcare education evidence synthesis. *BMC Medicine*. 2014;12(1).
95. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research and Policy*. 2005;10(1):45-53.
96. Rumrill PD, Fitzgerald SM, Merchant WR. Using scoping literature reviews as a means of understanding and interpreting existing literature. *Work (Reading, Mass)*. 2009;35(3):399-404.
97. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*. 1981;29(2):75-91.
98. Shenton AK, Hayter S. Strategies for gaining access to organisations and informants in qualitative studies. *Education for Information*. 2004;22(3-4):223-31.
99. Messick S. Validity. In RL Linn (Ed.), *Educational measurement* (pp. 13-103). Washington, DC: American Council on Education and National Council on Measurement in Education. 1989.
100. Mosier CI. A critical examination of the concepts of face validity. *Educational and Psychological Measurement*. 1947.
101. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Medical education*. 2004;38(3):327-33.
102. Schifferdecker KE, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. *Medical education*. 2009;43(7):637-44.
103. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. *Medical education*. 2007;41(8):737-45.
104. Stefanidis D, Heniford BT. The formula for a successful laparoscopic skills curriculum. *Archives of surgery* 2009;144:77-82.

105. Devine LA, Donkers J, Brydges R, Perelman V, Cavalcanti RB, Issenberg SB. An Equivalence Trial Comparing Instructor-Regulated With Directed Self-Regulated Mastery Learning of Advanced Cardiac Life Support Skills. *Simulation in Healthcare*. 2015;10(4):202-9.
106. Lineberry M, Park YS, Cook DA, Yudkowsky R. Making the case for mastery learning assessments: Key issues in validation and justification. *Academic Medicine*. 2015;90(11):1445-50.
107. McGaghie WC, Issenberg SB, Barsuk JH, Wayne DB. A critical review of simulation-based mastery learning with translational outcomes. *Medical education*. 2014;48(4):375-85.
108. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Medical education featuring mastery learning with deliberate practice can lead to better health for individuals and populations. *Academic Medicine*. 2011;86(11):e8-e9.
109. Yudkowsky R, Park YS, Lineberry M, Knox A, Ritter EM. Setting Mastery Learning Standards. *Academic medicine : journal of the Association of American Medical Colleges*. 2015;90(11):1495-500.
110. Hatala R, Cook DA, Brydges R, Hawkins R. Constructing a validity argument for the Objective Structured Assessment of Technical Skills (OSATS): A systematic review of validity evidence. *Advances in Health Sciences Education*. 2015;20(5):1149-75.
111. Muijtjens A, Hoogenboom R, Verwijnen G, Van der Vleuten C. Relative or absolute standards in assessing medical knowledge using progress tests. *Advances in Health Sciences Education*. 1998;3(2):81-7.
112. McLachlan JC. The relationship between assessment and learning. 2006.
113. van Empel PJ, Verdam MG, Strypet M, van Rijssen LB, Huirne JA, Scheele F, et al. Voluntary autonomous simulator based training in minimally invasive surgery, residents' compliance and reflection. *Journal of surgical education*. 2012;69(4):564-70.
114. Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM. Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. *Surgical endoscopy*. 2008;22(8):1887-93.
115. Bjerrum F, Maagaard M, Led Sorensen J, Rifbjerg Larsen C, Ringsted C, Winkel P, et al. Effect of Instructor feedback on skills retention after laparoscopic simulator training: Follow-up of a randomized trial. *Journal of surgical education*. 2014;72:53-60.

116. Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: A randomized trial. 2013;257:839-44.
117. Chiviacowsky S, Wulf G. Self-controlled feedback is effective if it is based on the learner's performance. Research quarterly for exercise and sport. 2005;76(1):42-8.
118. Sørensen SMD, Mahmood O, Konge L, Thinggaard E, Bjerrum F. Laser visual guidance versus two-dimensional vision in laparoscopy – A randomized trial. Surgical endoscopy. 2016;1-7.

10. PAPERS

Paper 1

Thinggaard E, Kleif J, Bjerrum F, Strandbygaard J, Gögenur I, Ritter E M, Konge L. Off-site training of laparoscopic skills, a scoping review using a thematic analysis. *Surgical Endoscopy* 2016;11:1-9.

Paper 2

Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *British Journal of Surgery* 2015;102:1106–1113.

Paper 3

Thinggaard E, Bjerrum F, Strandbygaard J, Gögenur I, Konge L. Ensuring competency of novice laparoscopic surgeons – Exploring standard setting methods and their consequences. *Journal of Surgical Education* 2016.

Paper 4

Thinggaard E, Bjerrum F, Strandbygaard J, Konge L, Gögenur I. Take-home training in laparoscopy facilitates distributed training, a randomized trial. *Surgery* 2016; Submitted.

Paper 5

Thinggaard E, Konge L, Bjerrum F, Strandbygaard J, Gögenur I, Spanager L. Take-home training in a simulation-based laparoscopy course. *Surgical Endoscopy* 2016.

Off-site training of laparoscopic skills, a scoping review using a thematic analysis

Ebbe Thinggaard^{1,2}  · Jakob Kleif³ · Flemming Bjerrum⁴ · Jeanett Strandbygaard⁴ · Ismail Gögenur¹ · E. Matthew Ritter⁵ · Lars Konge²

Received: 7 September 2015 / Accepted: 15 February 2016
© Springer Science+Business Media New York 2016

Abstract

Background The focus of research in simulation-based laparoscopic training has changed from examining whether simulation training works to examining how best to implement it. In laparoscopic skills training, portable and affordable box trainers allow for off-site training. Training outside simulation centers and hospitals can increase access to training, but also poses new challenges to implementation. This review aims to guide implementation of off-site training of laparoscopic skills by critically reviewing the existing literature.

Methods An iterative systematic search was carried out in MEDLINE, EMBASE, ERIC, Scopus, and PsychINFO, following a scoping review methodology. The included literature was analyzed iteratively using a thematic analysis approach. The study was reported in accordance with the

STructured apprOach to the Reporting In healthcare education of Evidence Synthesis statement.

Results From the search, 22 records were identified and included for analysis. A thematic analysis revealed the themes: access to training, protected training time, distribution of training, goal setting and testing, task design, and unsupervised training. The identified themes were based on learning theories including proficiency-based learning, deliberate practice, and self-regulated learning.

Conclusions Methods of instructional design vary widely in off-site training of laparoscopic skills. Implementation can be facilitated by organizing courses and training curricula following sound education theories such as proficiency-based learning and deliberate practice. Directed self-regulated learning has the potential to improve off-site laparoscopic skills training; however, further studies are needed to demonstrate the effect of this type of instructional design.

Electronic supplementary material The online version of this article (doi:[10.1007/s00464-016-4834-9](https://doi.org/10.1007/s00464-016-4834-9)) contains supplementary material, which is available to authorized users.

✉ Ebbe Thinggaard
ebbe.thinggaard@gmail.com

Jakob Kleif
jakobkleif@icloud.com

Flemming Bjerrum
fbjerrum@gmail.com

Jeanett Strandbygaard
jeanett78@gmail.com

Ismail Gögenur
ig@dadlnet.dk

E. Matthew Ritter
eritter@usuhs.edu

Lars Konge
lars.konge@regionh.dk

¹ Department of Surgery, Zealand University Hospital, Koege, Denmark

² Copenhagen Academy for Medical Education and Simulation, The Capital Region of Denmark, Blegdamsvej 9, Copenhagen 2300, Denmark

³ Department of Surgery, Hilleroed Hospital, Hilleroed, Denmark

⁴ Department of Obstetrics and Gynecology, Juliane Marie Centre, Rigshospitalet, Copenhagen, Denmark

⁵ Department of Surgery and NCR Simulation Consortium, Uniformed Services University / Walter Reed National Military Medical Center, Bethesda, MD, USA

Keywords Laparoscopy · Medical education · Simulation · Box trainer · Off-site · Technical skills

Simulation training has become a common method of acquiring the necessary skills for laparoscopic surgery. It can prepare surgical trainees and improve their skills before they are asked to operate on real patients. In some places, simulation training is now a requirement for certification as a specialist [1]. Laparoscopic simulation training can be done using simple box trainers (BTs) or virtual reality simulators (VRSs). A previous systematic review, which looked at the effect of simulation training in laparoscopic surgery, found that both VRS and BT provide effective training although BTs tend to be favored by trainees [2]. Two recent Cochrane reviews have looked at the literature regarding the use of BT in laparoscopic training. One of the reviews looked at the effect of training on trainees with no prior experience [3], and the other review looked at the effect of training on trainees with limited experience [4]. The conclusion was that box training is useful for acquiring laparoscopic skills. The evidence is strongest for trainees with no prior experience compared with trainees with limited experience. BTs are affordable and portable, giving trainees the freedom to train at their convenience [5]. Additionally, training on your own time would off load some training from increasingly restrictive duty hour requirements for surgical trainees. Despite these potential advantages, off-site training poses new challenges. Training unsupervised can lead to learning incorrect techniques and becoming overconfident in one's own abilities.

To help guide further implementation of off-site training, we need to critically evaluate the literature. The purpose of this review is to look at the literature regarding off-site training in laparoscopic surgery and to explore the instructional designs used in off-site training. As research in medical education is moving from demonstrating the effect of training on skills acquisition, to exploring the 'how,' 'why,' 'when,' and 'for whom' [6], we need to focus on how we implement training in our surgical specialty curricula.

Materials and methods

We performed a scoping review following the methodology described by Arksey and O'Mally [7] and further developed by Levac et al. [8]. In reporting the review, we followed the principles laid out in the STructured apprOach to the Reporting In healthcare education of Evidence Synthesis (STORIES) Statement [6]. The research group identified inclusion criteria as records exploring the use of BTs in off-site training of laparoscopic skills. We defined

off-site training as training taking place outside the location of regular instructional activities, such as simulation centers. We defined BTs as simple portable simulators, which can be used to train laparoscopic skills and are available at affordable prices. Records, which only looked at VRS, compared VRS to BTs, or validation studies of BTs were excluded from the study. Keywords were identified and a search string created (Table 1). The research group participated in the identification of the relevant literature from a variety of sources. A systematic search for the literature was carried out in MEDLINE, EMBASE, ERIC, Scopus, and PsychINFO. Records were included up to March 31, 2015, and were limited to the literature published in English. Two researchers worked independently searching the databases and then screened records by title/abstract, and finally, full texts articles were read. In case a disagreement arose on inclusion of a record, it was discussed until consensus could be reached. For conference abstracts, authors were contacted to find full text articles. The literature search also included a 'snowballing search,' looking through references from the included literature and for articles citing included literature. Scopus was used for identifying records from these sources. The literature was also identified by searching relevant Web sites such as the Fundamentals of Laparoscopic Surgery (FLS) and Society of Gastrointestinal and Endoscopic Surgeons (SAGES). Relevant reviews were also obtained, and their references searched for additional studies. The records were then analyzed, data were retrieved, and themes were identified in the literature using a thematic analysis approach [9].

Results

From the various database searches, 1978 records were identified. An additional 67 records were included from other sources. After removing duplicates, 1858 records

Table 1 Identification of key words and search string

Laparoscopy	Box trainers	Training
Laparoscopy	Boxtrainer	Training
Laparoscopic	Box-trainer (Box trainer)	Education
	Trainer	Simulation
	Blackbox	Practice
	Black-box	Skills
	Simulator	

Search String

(Laparoscopy OR Laparoscopic) AND (Simulator OR Boxtrainer OR Box-trainer OR Trainer OR Blackbox OR Black-box) AND (Training OR Education OR Simulation OR Skills OR Practice)*

*Adjusted for use in different databases

remained. Screening and reading full text articles resulted in 22 articles, which were included for analysis (Fig. 1).

Study characteristics

The included studies varied considerably, regarding the number of participants. There was a range from using a single surgeon's experience to up to 80 participants. Fourteen studies included only surgical trainees, four studies used medical students, two studies used mixed participants, and two studies used a single experienced surgeon. From the 22 studies, ten were comparative studies of which eight were randomized. Two of the randomized studies were blinded. Ten studies were descriptive, one study was a survey, and one study included training logs, a questionnaire, and an in-depth survey. Computerized rating was used in three studies. Two of these studies used metrics measured on VRS, and one study used the 'Imperial College Surgical Assessment Device (ICSAD),' a motion analysis tool used in BTs [10].

Reporting of records also varied, 12 articles were published in peer-reviewed journals, and ten were presented as abstracts at conferences and meetings (Table 2).

Effect of off-site training

The studies included in this review all reported that off-site training was a feasible training method; nevertheless, four of these made recommendations on improvements to training program in order for off-site training to become effective. Fifteen of the studies found that there was an effect of training off-site; however, four of these found that training on-site resulted in better training outcomes than training off-site. Five studies reported that participants underused off-site training.

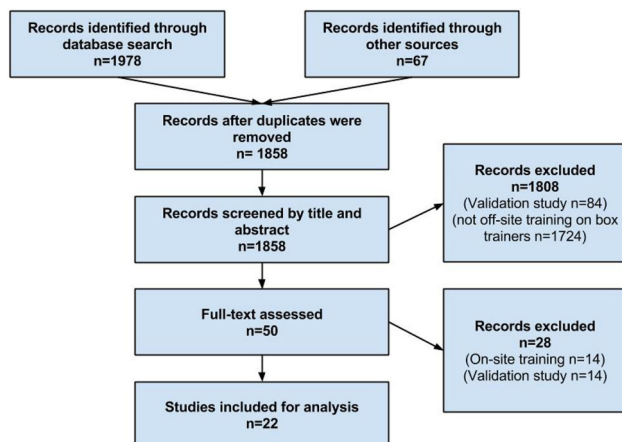


Fig. 1 Study identification, screening, and selection

Thematic analysis

Themes regarding instructional design were identified through a thematic analysis of the literature [9]. An example of the process is shown in Table 3. Themes included access to training, protected training time, distribution of training, goal setting and testing, task design, and unsupervised training. Learning theories underpinning the identified themes included proficiency-based learning, deliberate practice, and self-regulated learning (Table 4).

Discussion

Off-site training was recommended in the majority of the literature; however, several studies reported that changes to current off-site training were necessary to reach the full potential in off-site training. Making training accessible and available to trainees was a main theme represented in all studies. Lack of protected time for practice, cost of simulators, and physical access to training facilities were identified as barriers to training. Training off-site on portable and affordable trainers can help overcome these as they are available at affordable prices, make training easy accessible, and allow for the trainees to choose when and where to train [11].

Table 2 Study characteristics

Characteristic	Level	No. of studies
Reporting format	Abstract for presentation	10
	Article in peer-reviewed journal	12
Study participants	Medical students (n = 72)	4
	Surgical trainees (n = 304)	14
	Consultants (n = 2)	2
	Mixed groups (n = 29)	2
Study type	Randomized controlled study	8
	Comparative study	2
	Descriptive study of a group	8
	One person descriptive study	2
	Survey	1
	Mixed method	1
Study design	Pre–posttesting, two groups	9
	Pre–posttesting, one group	4
	Posttesting, two groups	1
	Posttesting, one group	1
	Observational	5
	Survey or mixed methods	2

Table 3 Examples of thematic analysis and underlying learning theories

Text from records:	Theme	Learning theory
'Performance scores were calculated and recorded for each task.' [20]	Goal setting	Proficiency-based learning
'..the box trainer is taken home for 6 weeks of autonomous training..' [23]	Unsupervised training	Self-regulated learning
'..there will be a more distributed training regimen..' [20]	Distribution of training	Deliberate practice Distributed practice

Distributed practice

Distribution of training has shown to be an effective instructional method [12]. When compared to massed training, distributed training is recommended in laparoscopic VRS training [13]. Traditionally, training of laparoscopic skills has been done on laparoscopic skills courses using massed training sessions. The availability of trainees, faculty, and, most importantly, non-portable simulation equipment has resulted in this often suboptimal training strategy. Portability in training makes it easier to distribute training sessions, but recommendation on how distributed training should be organized varies greatly. The studies included in this review recommend training schedules ranging from two times 30 min per week to 1 h per day [14–17]. Studies, which asked participants to log their training, also described a wide variation, from not using the BT at all, up to training 114 min per week [18–20]. The duration of training varied from 15 days to eight months. One study which looked at the optimal training schedule found that one-and-a-half-hour sessions once daily in the beginning of a training curriculum were effective in learning how to perform intra-corporal suturing [21]. Another study looking at the ideal training intervals in laparoscopic surgery found that short intervals were superior to having long intervals [22]. Studies demonstrate, however, that trainees do not train as much as recommended, indicating that a recommended training schedule on its own is insufficient in getting trainees to practice regularly [18, 20, 23]. Setting clear goals that the trainees have to reach has been described as a solution to this problem [18, 23]. Giving both clear recommendations on training intervals and having a goal to reach could help facilitate distributed practice in off-site training.

Proficiency-based learning

Maintaining participants interested and motivated in training regularly can be a challenge. Implementing tests can help provide goals for trainees. Testing alone has a positive effect on retention of learned skills [24]. When implementing tests however, it is important that they are supported by evidence of validity [25], that the pass/fail

level is fair, and established using a evidence-based standard setting method [26]. Proficiency-based standard setting has been implemented using performance levels of experienced laparoscopic surgeons. A proficiency-based learning strategy has been used in laparoscopic simulation training on VRS and shown to improve OR performance [27]. A similar effect has been demonstrated when training on BT [28]. However, assessing performance on tasks in a BT requires either a system for self-monitoring, a method of automated scoring, or the feedback from an experienced laparoscopic surgeon. Proficiency-based training is recommended [13], and programs using home training on BT should set criteria for proficiency [24].

Protected time for training

Training off-site blurs the boundaries between work time and private time. Popularly, this issue could be described as 'time is money' and training is to some trainees seen as a part of work for which economic compensation is expected. Acquiring laparoscopic skills requires training, and training during duty hours is often interrupted by clinical duties and limited by duty-hours restrictions. Protected time is mentioned as a solution to overcome this problem. A study looking at VRS training described this as an important factor in motivating trainees to complete a laparoscopic skills curriculum [29]. However, this may vary greatly across different countries. In some countries, training is seen as part of one's work responsibilities, and in other countries training is expected to be done outside duty hours. Whether training is part of the duty hours or not, off-site training provides easier access to training and can be applied in both settings.

Deliberate practice

Feedback provided by an experienced surgeon has been a natural part of training surgical skills used in Halstead's principles of apprenticeship training, where the trainees learn from the master. Exploring how experts achieve their skills has led to the educational theory of deliberate practice. Deliberate practice uses immediate feedback,

Table 4 Overview of included studies and themes

Author	Themes	Study description
Gue [14]	Access to training Distribution of training	Study describing a single experienced surgeon training off-site
Robinson et al. [46]	Access to training Distribution of training Competition	Study comparing on-site training on a Tower trainer versus off-site training on a mirrored box trainer. Participants trained on basic tasks
Griffin et al. [19]	Access to training Task design	Study comparing off-site training versus no training. Participants trained suturing
Arden et al. [47]	Access to training Goal setting Task design	Study comparing off-site training versus no training. Participants trained procedure-like tasks
Morgan et al. [33]	Access to training Task design	Study describing off-site training using one group. Participants practiced basic tasks, suturing and advanced tasks
Morgan et al. [32]	Access to training Task design	Study describing off-site training combined with training on-site using one group. Participants trained increasingly difficult tasks
Stovall et al. [48]	Access to training Distribution of training	Study describing off-site training using one group. Participants trained basic tasks
Okrainec et al. [35]	Access to training Distribution of training Unsupervised training Goal setting	Study comparing off-site training using telesimulation versus off-site training with no telesimulation. Participants trained using the FLS tasks
Russo et al. [15]	Access to training	Study describing off-site training using one group. Participants trained using the FLS peg-transfer task
Rabie [49]	Access to training Distribution of training	Study describing a single experienced surgeon training off-site
Russo et al. [16]	Access to training Goal setting Task design	Study describing off-site training using one group. Participants trained using basic tasks
Kobayashi et al. [50]	Access to training Goal setting	Study describing off-site training using one group. Participants trained using the FLS tasks
Harrity et al. [51]	Access to training Distribution of training	Study comparing off-site training versus no training. Participants trained basic tasks and suturing
Chummun et al. [34]	Access to training Distribution of training Goal setting Task design Unsupervised training	Study describing off-site training combined with on-site training, using one group. Participants trained on basic tasks, suturing and advanced tasks
Harrity et al. [52]	Access to training Distribution of training	Study describing off-site training using one group. Participants trained suturing
Nakamura et al. [31]	Unsupervised training Task design Distribution of training	Study comparing off-site training versus no training. Participants trained basic tasks
Arthur et al. [53]	Access to training	Study comparing off-site training versus no training. Participants trained basic tasks
Korndorffer et al. [20]	Access to training Protected training time Distribution of training Goal setting Unsupervised training	Study comparing off-site training versus on-site training. Participants trained using the FLS peg-transfer and suturing task

Table 4 continued

Author	Themes	Study description
van Empel et al. [23]	Access to training Protected training time Distribution of training Goal setting Unsupervised training	A study exploring compliance and reflection of trainees using training logs, a questionnaire and an in-depth survey
Harvey et al. [17]	Access to training Unsupervised training	Study comparing off-site training versus no training. Participants trained suturing
Caban et al. [54]	Access to training	Study comparing the combined effect of off-site training and on-site training versus only on-site training. Participants trained using the four FLS tasks and basic tasks
Zapf et al. [18]	Access to training Goal setting Distribution of training Compulsory training	A survey study exploring the use of off-site training

distributed training, and well-defined tasks [30]. Deliberate practice was used as a base for one of the included studies which showed that off-site training leads to more random training schedule which could increase retention of skills [20]. Well-defined tasks are also part of the principles of deliberate practice. Tasks for laparoscopic training can be designed in a variety of ways ranging from basic tasks and suturing to tasks resembling surgical procedures. Several studies used one task, primarily suturing, and others included several tasks [31]. One study, however, used increasingly difficult tasks throughout a 6-month laparoscopic training course [32]. Using increasingly complex tasks requires new instructions, which was provided by faculty through on-site training sessions. Principles of deliberate practice using appropriate task design can be implemented in off-site training, although providing feedback remains a challenge.

Proctored training and feedback

Proctored training where supervision is provided from an experienced laparoscopic surgeon is difficult to integrate in off-site training. This type of instruction is usually only accessible during training sessions taking place on-site. A combination of off-site and on-site training can be used and has shown to improve laparoscopic skills training on BT [32–34]. Using on-site training, however, means increased resources spent on training, as faculty attendance is required for proctored training sessions. Administration costs will also increase, as training sessions need to be planned. Proctored training usually requires the physical presence of faculty at the training site. One novel study, however, explored the use of telesimulation in laparoscopic skills training. Proctored telesimulation was shown to be an

effective method of instruction and resulted in a passing rate of 100 % on the Fundamentals of Laparoscopic Surgery (FLS) test [35]. Feedback during proctored training can vary substantially though. Objective rating scales for assessment, such as OSATS [36] and GOALS [37], can help structure this type of feedback.

Feedback is an integral part of deliberate practice and can be provided in different ways. Feedback is either qualitative or quantitative. A previous study looked at the different types of feedback, qualitative (comments), quantitative (marks), and mixed qualitative and quantitative feedback among fifth and sixth grade school children. In this study, an effect of qualitative feedback gave better results, and mixed feedback did not seem to be superior [38]. The use of different types of feedback has been explored in endotracheal intubation. A positive effect of quantitative feedback was seen on short-term evaluation, but no effect was demonstrated when looking at retention [39]. In laparoscopic training, the effect of structured quantitative feedback has been shown to be beneficial when training complex procedural skills [40]. When practicing basic laparoscopic skills however, a proctored training system with feedback was not superior [41]. Although feedback is usually provided by experienced faculty, a different approach is the use of self-recording. Self-recording using a systematic scoring system relying on time and number of errors provides the trainees with quantitative feedback. A study providing participants with the means of rating their own performance and having a well-defined proficiency level lead to 100 % of participants passing the FLS Manual Skill Exam [42, 43]. Training unsupervised using self-monitoring of performance can be beneficial as trainees have more responsibility of their own training, leading to a more autonomous approach to training.

Self-regulated learning

Having trainees organize their own training, choosing when to practice and in which order tasks are used, stems from the educational theories of self-regulated learning (SRL). This educational approach is recommended in a recent systematic review of SRL in simulation-based training [44]. A new approach to self-regulated learning called directed self-regulated learning (DSRL) is emerging. DSRL has in one study proven effective in simulation training for lumbar puncture when looking at retention of skills [45]. For off-site training in laparoscopy, DSRL could be well suited as it may help overcome some of the problems unsupervised training poses. Training off-site can lead to trainees using the wrong technique or failing to reach a set proficiency level. Trainees can be instructed prior to unsupervised training by faculty or throughout the off-site training by the use of instructional video clips. DSRL shows great promise, as an instructional method for off-site training. Studies should be conducted to explore this approach in acquiring laparoscopic techniques.

Recommendations

Through this review, several themes have been explored and educational theories for optimal instructional design discussed. Based on our finding, we would recommend that when implementing off-site training in laparoscopy, accessible box trainers that are portable and affordable to the trainees should be used. The instructional design should be based on the principles of deliberate practice and proficiency-based learning. Courses should be structured to ensure that participants practice several days a week during the training program. Curricula should use tasks with a set proficiency level to reach, have clear learning objectives and increasing difficulty. On-site training should be incorporated as part of the training course if feasible. Feedback should be provided during off-site training. Feedback can be implemented in off-site training using assessment tools with evidence of validity, either through on-site proctored training sessions, telesimulation, or by allowing trainees to self-monitor their performance.

Randomized trials examining the effect of off-site training are needed to demonstrate whether this off-site training is effective when compared with regular training facilities. Many factors like motivation, logistics and whether training is mandatory can influence the success of off-site training as a training strategy. Simply equipping trainees with a BT and relying on them to practice without an overall training structure is an antiquated strategy and should be abandoned.

Off-site training of laparoscopic skills on a BT can be an effective method of instruction. Training on BT off-site can be improved by following the principles of proficiency-based learning and deliberate practice. Directed self-regulated learning has the potential to improve off-site laparoscopic skills training; however, further studies are needed to demonstrate the effect of this type of instructional design.

Acknowledgements We would like to thank the staff at the Hospital Research Library, Rigshospitalet, for their help and support in conducting this review.

Compliance with ethical standards

Disclosures Dr. Ebbe Thinggaard reports non-restricted educational grant from Johnson and Johnson. Dr. E. Matthew Ritter reports educational grant support from the Henry M. Jackson foundation for the advancement of military medicine. Drs. Jakob Kleif, Flemming Bjerrum, Jeanett Strandbygaard, Ismail Gögenur, and Lars Konge have no conflicts of interest or financial ties to disclose.

References

1. Soper NJ, Fried GM (2008) The fundamentals of laparoscopic surgery: its time has come. *Bull Am Coll Surg* 93(9):30–32
2. Zendejas B, Brydges R, Hamstra SJ, Cook DA (2013) State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 257(4):586–593
3. Nagendran M, Toon C, Davidson B, Gurusamy K (2014) Laparoscopic surgical box model training for surgical trainees with no prior laparoscopic experience. *Cochrane Database Syst Rev*. doi:10.1002/14651858.CD010479.pub2
4. Gurusamy KS, Nagendran M, Toon CD, Davidson BR (2014) Laparoscopic surgical box model training for surgical trainees with limited prior laparoscopic experience. *Cochrane Database Syst Rev*. doi:10.1002/14651858.CD010478.pub2
5. Thinggaard E, Bjerrum F, Strandbygaard J, Gogenur I, Konge L (2015) Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *Br J Surg* 102(9):1106–1113
6. Gordon M, Gibbs T (2014) STORIES statement: publication standards for healthcare education evidence synthesis. doi:10.1186/s12916-014-0143-0
7. Arksey H, O'Malley L (2005) Scoping studies: towards a methodological framework. *Int J Soc Res Methodol Theory Pract* 8(1):19–32
8. Levac D, Colquhoun H, O'Brien KK (2010) Scoping studies: advancing the methodology. *Implement Sci*. doi:10.1186/1748-5908-5-69
9. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A (2005) Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy* 10(1):45–53
10. Moorthy K, Munz Y, Dosis A, Bello F, Darzi A (2003) Motion analysis in the training and assessment of minimally invasive surgery. *Minim Invasive Ther Allied Technol* 12(3):137–142
11. Partridge R, Hughes M, Brennan P, Hennessey I (2015) There is a worldwide shortfall of simulation platforms for minimally invasive surgery. *J Surg Simul*. doi:10.1102/2051-7726.2015.0003
12. Benjamin AS, Tullis J (2010) What makes distributed practice effective? *Cogn Psychol* 61(3):228–247

13. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 241(2):364–372
14. Gue S (1995) Home-made videoscopic trainer for operative laparoscopic surgery. *Aust N Z J Surg* 65:820–821
15. Russo MA, Tsuda S (2010) Portable, self-practice laparoscopic box trainers underutilized by surgical trainees. In Paper presented at the 2010 scientific session of the Society of American Gastrointestinal and Endoscopic Surgeons, SAGES 2010, National Harbor, MD United States, 14 April 2010
16. Russo M, Tsuda S (2011) Assessment of a mandatory curriculum for laparoscopic skills using a portable, self-practice box trainer outside of duty hours. In Paper presented at the 2011 Scientific Session of the Society of American Gastrointestinal and Endoscopic Surgeons, SAGES San Antonio, TX United States, 30 March 2011
17. Harvey LFB, King L, Hur HC (2013) Challenges associated with a self-contained simulation curriculum using a home laparoscopic skills trainer. In Paper presented at the 42nd Global Congress of Minimally Invasive Gynecology AAGL 2013, Washington, DC United States, 10 November 2013
18. Zapf M, Ujiki M (2013) Surgical resident evaluations of portable laparoscopic box trainers. *Surg Innov*. doi:10.1177/1553350614535858
19. Griffin S, Kumar A, Burgess N, Donaldson P (2006) Development of laparoscopic suturing skills: a prospective trial. *J Endourol* 20(2):144–148
20. Korndorffer JR Jr, Bellows CF, Tekian A, Harris IB, Downing SM (2012) Effective home laparoscopic simulation training: a preliminary evaluation of an improved training paradigm. *Am J Surg* 203(1):1–7
21. De Win G, Van Bruwaene S, De Ridder D, Miserez M (2013) The optimal frequency of endoscopic skill labs for training and skill retention on suturing: a randomized controlled trial. *J Surg Educ* 70(3):384–393
22. Stefanidis D, Walters KC, Mostafavi A, Heniford BT (2009) What is the ideal interval between training sessions during proficiency-based laparoscopic simulator training? *Am J Surg* 197(1):126–129
23. van Empel PJ, Verdam MG, Strypet M, van Rijssen LB, Huijse JA, Scheele F, Bonjer HJ, Meijerink WJ (2012) Voluntary autonomous simulator based training in minimally invasive surgery, residents' compliance and reflection. *J Surg Educ* 69(4):564–570
24. Kromann CB, Jensen ML, Ringsted C (2009) The effect of testing on skills learning. *Med Educ* 43(1):21–27
25. Downing SM (2003) Validity: on meaningful interpretation of assessment data. *Med Educ* 37(9):830–837
26. Norcini JJ, Shea JA (1997) The credibility and comparability of standards. *Appl Meas Educ* 10(1):39–59
27. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 236(4):458–463 (**discussion 463–454**)
28. Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayed R, Fried GM (2010) Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room—a randomized controlled trial. *Am J Surg* 199(1):115–120
29. Strandbygaard J, Bjerrum F, Maagaard M, Riffbjerg Larsen C, Ottesen B, Sorensen JL (2014) A structured four-step curriculum in basic laparoscopy: development and validation. *Acta Obstet Gynecol Scand* 93(4):359–366
30. Ericsson KA (2004) Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 79(10 Suppl):S70–S81
31. Nakamura LY, Martin GL, Fox JC, Andrews PE, Humphreys M, Castle EP (2012) Comparing the portable laparoscopic trainer with a standardized trainer in surgically naïve subjects. *J Endourol* 26(1):67–72
32. Morgan D, O'Sullivan R, Prendiville W (2010) A study of a novel 'take-home' Laparoscopic simulator and program. In Paper presented at the 39th Global Congress of Minimally Invasive Gynecology, AAGL 2010, Las Vegas, NV United States, 11 August 2010
33. Morgan D, Prendiville W, O'Sullivan R (2010) A laparoscopic surgical skills program utilising a 'take-home' laparoscopic simulator. In Paper presented at the 39th Global Congress of Minimally Invasive Gynecology, AAGL 2010, Las Vegas, NV United States, 08 November 2010
34. Chummun K, Burke JP, O'Sullivan R, Prendiville W (2012) The influence of a 'take home' box trainer on laparoscopic performance for gynaecological surgeons. *Gynecol Surg* 9:303–308
35. Okrainec A, Henao O, Azzie G (2010) Telesimulation: an effective method for teaching the fundamentals of laparoscopic surgery in resource-restricted countries. *Surg Endosc* 24(2):417–422
36. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84(2):273–278
37. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondre K, Stanbridge D, Fried GM (2005) A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 190(1):107–113
38. Butler R (1987) Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest, and performance. *J Educ Psychol* 79(4):474–482
39. Manzone J, Tremblay L, You-Ten KE, Desai D, Brydges R (2014) Task- versus ego-oriented feedback delivered as numbers or comments during intubation training. *Med Educ* 48(4):430–440
40. Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, Gluud C, Grantcharov T, Ottesen B, Sorensen JL (2013) Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Ann Surg* 257(5):839–844
41. Snyder CW, Vandromme MJ, Tyra SL, Hawn MT (2009) Proficiency-based laparoscopic and endoscopic training with virtual reality simulators: a comparison of proctored and independent approaches. *J Surg Educ* 66(4):201–207
42. Ritter EM, Scott DJ (2007) Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery. *Surg Innov* 14(2):107–112
43. Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM (2008) Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. *Surg Endosc* 22(8):1887–1893
44. Brydges R, Manzone J, Shanks D, Hatala R, Hamstra SJ, Zendejas B, Cook DA (2015) Self-regulated learning in simulation-based training: a systematic review and meta-analysis. *Med Educ* 49(4):368–378
45. Brydges R, Nair P, Ma I, Shanks D, Hatala R (2012) Directed self-regulated learning versus instructor-regulated learning in simulation training. *Med Educ* 46(7):648–656
46. Robinson JK, Kushner DM (2006) Development and validation of a home-based, mirrored, gynecologic laparoscopy trainer. *J Minim Invasive Gynecol* 13(2):102–107
47. Arden D, Hacker MR, Jones DB, Awtrey CS (2008) Description and validation of the Pelv-Sim: a training model designed to improve gynecologic minimally invasive suturing skills. *J Minim Invasive Gynecol* 15:707–711

48. Stovall DW, Redick DL, Prettyman ST (2010) The usefulness of a personal laparoscopic trainer in resident education. In Paper presented at the 39th Global Congress of Minimally Invasive Gynecology, AAGL 2010, Las Vegas, NV United States, 08 November 2010
49. Rabie ME (2010) Acquiring laparoscopic suturing skills using a homemade trainer. 42 (Rabie) Faculty of Medicine, Jazan University, Jazan, Saudi Arabia, pp 149–151
50. Kobayashi SA, Jamshidi R, O'Sullivan P, Palmer B, Hirose S, Stewart L, Kim EH (2011) Bringing the skills laboratory home: an affordable webcam-based personal trainer for developing laparoscopic skills. *J Surg Educ* 68:105–109
51. Harrity C, Prendiville W, Ahmed S, Reus E, O'Sullivan R (2011) The benefits of a take-home trainer on the development of laparoscopic surgical skills. In Paper presented at the 20th Annual Congress ESGE, London, United Kingdom, 21 September 2011
52. Harrity C, O'Sullivan R, Prendiville W (2012) Effects of a home-based simulation programme. In Paper presented at the 21st Annual Congress ESGE, Paris France, 14 September 2012
53. Arthur LM, Moug SJ, Gough V, Vella M, Renwick A (2012) Laparoscopic training at home: a pilot study. In Paper presented at the 19th international congress of the European Association for Endoscopic Surgery, EAES 2011, Torino, Italy, 15 June 2011
54. Caban AM, Guido C, Silver M, Rossidis G, Sarosi G, Ben-David K (2013) Use of collapsible box trainer as a module for Resident Education. *J Soc Laparoendo* 17:440–444

Validity of a cross-specialty test in basic laparoscopic techniques (TABLT)

E. Thinggaard^{1,3}, F. Bjerrum², J. Strandbygaard², I. Gögenur³ and L. Konge¹

¹Centre for Clinical Education, Capital Region of Denmark, and ²Juliane Marie Centre, Rigshospitalet, Copenhagen and ³Department of General Surgery, Roskilde and Koege Hospital, Koege, Denmark

Correspondence to: Dr E. Thinggaard, Department of General Surgery, Roskilde and Koege Hospital, Lykkebaekvej 1, 4600 Koege, Denmark (e-mail: ebbe.thinggaard@gmail.com)

Background: Box trainer systems have been developed that include advanced skills such as suturing. There is still a need for a portable, cheap training and testing system for basic laparoscopic techniques that can be used across different specialties before performing supervised surgery on patients. The aim of this study was to establish validity evidence for the Training and Assessment of Basic Laparoscopic Techniques (TABLT) test, a tablet-based training system.

Methods: Laparoscopic surgeons and trainees were recruited from departments of general surgery, gynaecology and urology. Participants included novice, intermediate and experienced surgeons. All participants performed the TABLT test. Performance scores were calculated based on time taken and errors made. Evidence of validity was explored using a contemporary framework of validity.

Results: Some 60 individuals participated. The TABLT was shown to be reliable, with an intra-class correlation coefficient of 0.99 ($P < 0.001$). ANOVA showed a difference between the groups with different level of experience ($P < 0.001$). The Bonferroni correction was used to confirm this finding. A Pearson's r value of 0.73 ($P < 0.001$) signified a good positive correlation between the level of laparoscopic experience and performance score. A reasonable pass–fail standard was established using contrasting groups methods.

Conclusion: TABLT can be used for the assessment of basic laparoscopic skills and can help novice surgical trainees in different specialties gain basic laparoscopic competencies.

Paper accepted 16 April 2015

Published online in Wiley Online Library (www.bjs.co.uk). DOI: 10.1002/bjs.9857

Introduction

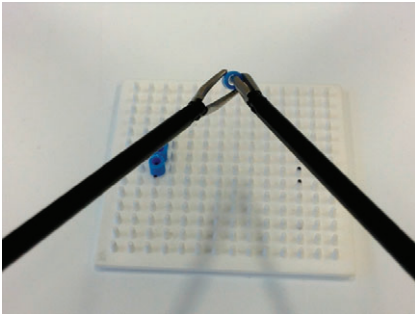
Minimally invasive techniques, and laparoscopy in particular, have become widespread in present clinical practice¹. Training is required to reach competency in laparoscopic skills and, because the level of surgical skill is related directly to the outcome of operation, recent research has highlighted the importance of competence². Laparoscopic surgery requires specific psychomotor skills as depth perception is missing, instruments are fixed at skin level and there is a limited range of movement³. To overcome this, laparoscopic techniques can be trained on box or virtual reality trainers outside the operating theatre. Training surgical skills using a simulator can shorten operating times, increase operative skills, and reduce the risk of both intra-operative and postoperative complications^{4–7}.

Criterion-based assessment and training has gained ground, and it is generally accepted that trainees should

train to reach a predefined level of proficiency^{8,9}. Assessment of skills ensures a relevant level of competency has been reached, and increases the motivation for trainees to practise^{8,10}. This level of competency, however, should be assessed using a test supported by evidence of validity^{11–14}. Until now, training systems for laparoscopy have been developed independently for each specialty (general surgery, urology and gynaecology), including learning advanced skills such as suturing. The Fundamentals of Laparoscopic Surgery (FLS) is the most widespread of these training systems, and is used for credentialing surgeons during specialty training¹⁵. However, there is still a need for a skills training and assessment tool for novice laparoscopic surgeons to use before performing supervised surgery on patients⁸.

The aims of this study were to explore evidence of validity for a test of basic laparoscopy skills, and to establish a reasonable pass–fail standard.

Task 1: coordination

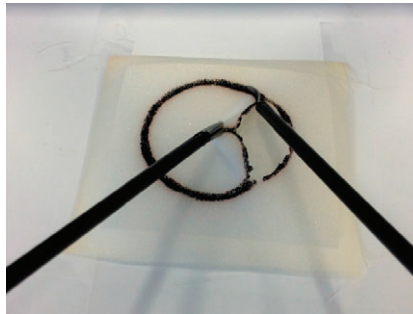


A bead to peg transfer task, in which four beads were transferred from pegs on one side of a pegboard to the other and back again by grasping the bead, transferring the bead from one instrument to another instrument, and then placing it back on the pegboard.

Errors were counted whenever a bead was dropped. If the bead fell outside reaching distance, this was counted as two errors. Each error resulted in 20s of penalty time.

The mean task score for a group of experienced surgeons was 427.

Task 2: cutting

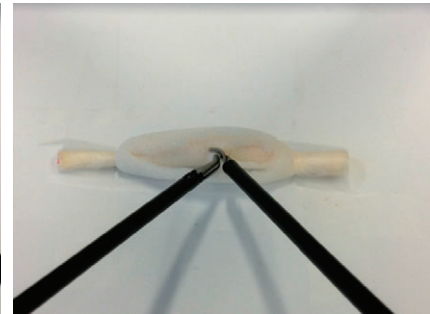


The purpose of this task was to cut a circle with a 2.5-cm radius from a washcloth, while cutting only within a 3-mm thick black line.

Errors were counted for each cut outside the line. Each error resulted in 20s of penalty time.

The mean task score for a group of experienced surgeons was 361.

Task 3: sharp dissection



The objective of this task was to dissect a rolled up washcloth (9 × 9cm), and to locate and view a vessel simulated by a balloon. The task was completed when two 2-mm black horizontal lines 2cm apart were fully exposed.

Errors were counted when damaging the 'vessel'. Each error resulted in 60s of penalty time.

The mean task score for a group of experienced surgeons was 489.

Task 4: blunt dissection

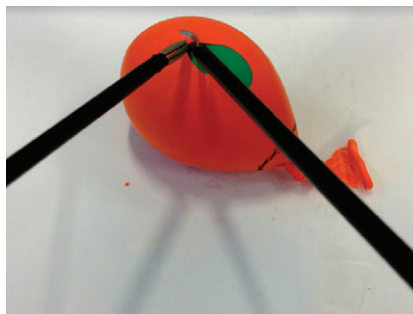


The aim of this task was to dissect a roll of cotton (5 × 9cm) using a blunt technique, and viewing a vessel simulated by a balloon. The task was completed when the 2-mm black lines 2cm apart were fully exposed.

Errors were counted when bits of cotton were torn from the roll as a result of the exercise. Errors resulted in 30s penalty time.

The mean task score for a group of experienced surgeons was 476.

Task 5: cyst removal

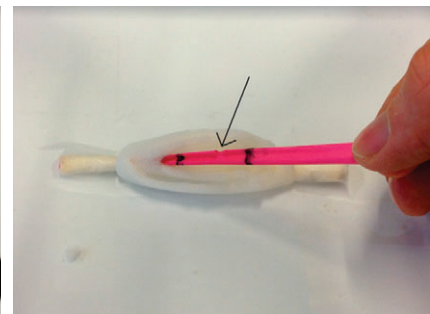


This task consisted of a balloon wrapped inside another balloon. The inner balloon was filled with 60ml ultrasound gel. The objective was to remove the inner balloon from the outer balloon by dissecting the outer balloon.

Errors were counted when the inner balloon was perforated. Errors resulted in 60s of penalty time.

The mean task score for a group of experienced surgeons was 393.

Example of an error



An error has been made in task 3, sharp dissection. The 'vessel' has been damaged.

Fig. 1 Description of tasks and errors



a TABLT training kit



b TABLT in use

Fig. 2 **a** Training and Assessment of Basic Laparoscopic Techniques (TABLT) training kit, including tasks. **b** Surgical trainee using TABLT

Methods

The study was submitted for evaluation to the regional ethics committee, which determined that no approval was needed (H-3-2013-FSP66).

Training and Assessment of Basic Laparoscopic Techniques

The Training and Assessment of Basic Laparoscopic Techniques (TABLT) test was developed during separate laparoscopy training programmes for general surgeons and gynaecologists. The tasks were developed with a focus on appropriate functional task alignment in order to enhance the transfer of learning¹⁶. Faculty and course participants provided feedback and adjustments were made to ensure relevance of the tasks. A pilot study was performed with two experienced surgeons and eight novices to ensure a reasonable level of difficulty and adjust the scoring system. Five tasks were included in the TABLT test, with its content reflecting basic laparoscopic techniques (Fig. 1). The tasks covered appropriate handling of laparoscopic instruments, cutting, blunt dissection and sharp dissection. The test also considered hand–eye coordination, guiding instruments via a screen, ambidexterity, accommodating the fulcrum effect and economy of movement.

Based on elements used in the FLS training ratings system¹⁷, a scoring system was developed taking account

Table 1 Participants

	Novice (n = 20)	Intermediate (n = 20)	Experienced (n = 20)	Total (n = 60)
Age (years)	24–31	27–41	31–58	24–58
Sex ratio (M : F)	7 : 13	11 : 9	13 : 7	31 : 29
Specialty				
Surgery	11	13	10	34
Urology	3	5	5	13
Gynaecology	6	2	5	13

of time and number of errors (Fig. 1). For each task, a score was calculated by subtracting the time spent on the task and a task-specific penalty score from a maximum time of 600 s, using the formula: task score = 600 – completion time – (no. of errors × penalty time per error).

The task scores were then standardized by dividing the task score by the mean score from a group of experienced surgeons and then multiplying it by 100. A performance score for the whole test was calculated as the sum of the five standardized task scores. The scoring system was tested during the pilot study.

Establishing validity evidence

A cohort of laparoscopic surgeons and surgical trainees were recruited and evidence of validity established using a contemporary framework of validity¹⁴. In accordance with

Table 2 Summary of validity evidence

Source of validity evidence	Questions related to each source of evidence	Validity evidence for TABLT
Content	Does the content reflect the underlying construct?	Tasks are aligned with the construct
Response process	Are sources of bias reduced?	Assessment can be done blinded, and calculation of the score automated
Internal structure	Is the test score reliable?	A high level of reliability shown: ICC = 0.99 ($P < 0.001$)
Relation to other variables	Does the test score correlate with a known measure of competence?	Novices, intermediates and experts score significantly differently ($P \leq 0.003$, ANOVA with Bonferroni correction) Test score correlates with operative experience: Pearson correlation $r = 0.73$ ($P < 0.001$).
Consequences of testing	What are the consequences of the pass–fail score?	Two of 20 of experts failed and two of 20 of novices passed the test

TABLT, Training and Assessment of Basic Laparoscopic Techniques; ICC, intraclass correlation coefficient.

this framework, evidence of validity was collected from five sources: content, response process, internal structure, relation to other variables and consequences of the test. Participants were recruited from three different specialties (general surgery, gynaecology and urology). Participants were recruited by e-mail through departmental heads, consultants responsible for training and direct contacts. They were divided into three groups according to level of laparoscopic experience. Novices had no previous experience in laparoscopic surgery, and less than 2 h of training on either a box trainer or virtual reality trainer; those with an intermediate level of experience had performed between one and 100 laparoscopic procedures; and experienced surgeons had carried out more than 100 laparoscopic procedures. Both intermediate and experienced surgeons were undertaking laparoscopic surgery in their current places of work. The TABLT test was administered on a portable tablet trainer¹⁸ (Fig. 2). Testing was performed after work or on days off, according to participant availability. All performed the TABLT test twice. The first attempt was to familiarize themselves with the set-up and scoring system. The second attempt was used for assessment purposes; it was rated on site by the corresponding author and afterwards by a blinded assessor using a video recording. The blinded assessor was another member of the research group who practised rating using videos recorded during a pilot study. Three videos, with examples of different types of error, were used.

Statistical analysis

Statistical analysis was performed to explore the internal structure. The intraclass correlation coefficient (ICC) was calculated, with single measures and absolute agreement definition. ANOVA was used to explore relationships with other variables. Differences between groups of laparoscopic surgeons with various levels of experience were analysed. A groupwise comparison

using ANOVA with Bonferroni correction was done to identify differences between groups for each of the pairings – novice *versus* intermediate, novice *versus* experienced, and intermediate *versus* experienced surgeons. The Pearson product–moment correlation coefficient was calculated to examine any correlation between the number of procedures performed and the test score. A Pearson's r value of 0.7 was considered an acceptable measure of correlation. $P < 0.050$ was considered statistically significant in the aforementioned tests. The contrasting groups method was used to set the pass–fail level, and the consequence of applying this was reported using relative frequencies converted to percentages. A pass–fail score that passed at least 85 per cent of the experienced surgeons and failed at least 85 per cent of the novices was considered reasonable. SPSS® version 20.0 (IBM, Armonk, New York, USA) was used for statistical analysis.

Results

All 60 laparoscopic surgeons and trainees (Table 1) completed the TABLT test twice. The second attempt was rated on site and by a blinded video assessor, resulting in 120 ratings.

Validity evidence is summarized in Table 2. The internal structure was explored by analysing test reliability. A high level of reliability was shown, with an ICC of 0.99 ($P < 0.001$). Relationships with other variables, examined by analysing variation in performance scores between novice, intermediate and experienced surgeons, are shown in Fig. 3. A significant difference between these groups was found ($P < 0.001$). The test discriminated between novices and experienced surgeons ($P < 0.001$), novice and intermediate surgeons ($P = 0.003$), and intermediate and experienced surgeons ($P < 0.001$). There was a correlation between the level of laparoscopic surgical experience and the test score, with a Pearson correlation r value of 0.73 ($P < 0.001$) (Fig. 4). A pass–fail level was established at

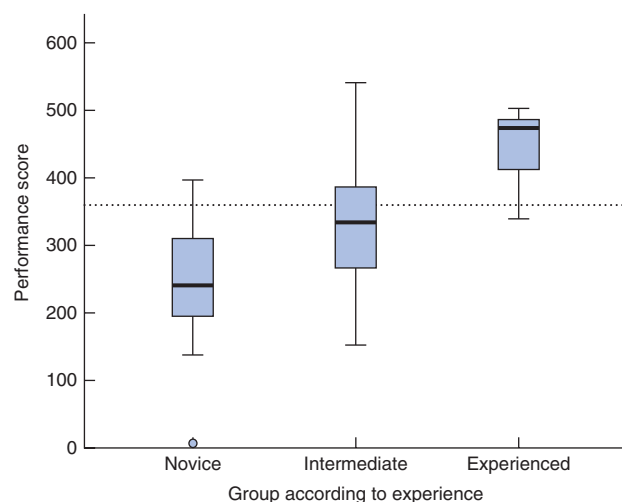


Fig. 3 Box plot of performance scores in relation to level of experience. Median values (horizontal lines), i.q.r. (boxes), and range (error bars) excluding outlier (circle) are shown. The dotted line indicates the pass–fail level. Mean(s.d.) scores for novice, intermediate and experienced surgeons were 244(88), 331(94) and 446(52) respectively

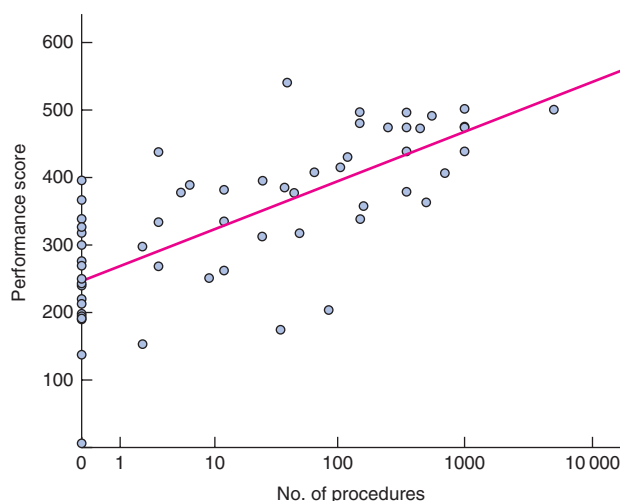


Fig. 4 Performance scores according to the level of experience expressed as number of procedures. Linear $R^2 = 0.526$

358 points using the contrasting groups methods (Fig. 5). The consequence of this pass–fail level was that two of 20 novices passed the test and two of 20 experienced surgeons failed it.

Discussion

The design of TABLT was inspired by another laparoscopic training and testing model, FLS¹⁹. The tasks included in

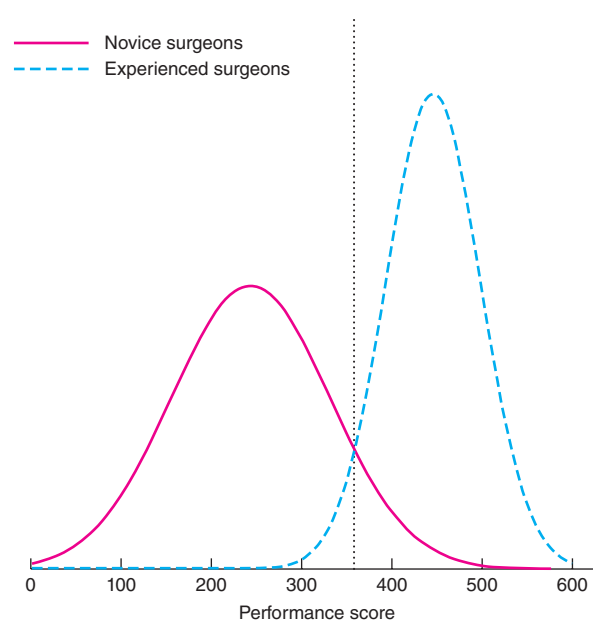


Fig. 5 Standard setting using contrasting groups method. The dotted line indicates the pass–fail level

TABLTL all reflected laparoscopic techniques that ensured a functional alignment of content with the construct, thereby providing evidence of validity from content. In contrast to the FLS, a laparoscopic suturing task was not included in TABLT, because this was considered a more advanced laparoscopic skill not performed by most novices. Camera navigation, on the other hand, is an essential skill and the focus of the Laparoscopic Skills Training and Testing used in gynaecology²⁰. Camera navigation is a relevant skill to train, but requires a movable camera, which is not currently included in the TABLT test. Other test systems of laparoscopic skills focus on basic movement and coordination skills²¹. TABLT includes these skills, but also surgical techniques such as cutting, blunt dissection and sharp dissection. Cutting and dissection are important laparoscopic techniques to master, especially as functional task alignment is high and it can help with instrument familiarization. The TABLT tasks reflect many of the skills needed when surgical trainees perform their first supervised laparoscopic operation on a patient.

Reducing sources of bias and ensuring that the intended response was elicited when administering the test provided evidence from the response process. Bias from data entry was countered by using simple spreadsheets that were preformatted to perform automatic score calculations. This also facilitated maintenance of data integrity, as data in spreadsheets are easily accessible and automated score calculations are transparent. Participants demonstrated an

appropriate response during testing because their first attempt allowed them to develop a strategy for doing the tasks, and they used this in the second attempt, which was the one used for rating.

The results showed a high ICC value, which demonstrated that the scoring system of the TABLT was reliable. A high level of reliability supports evidence of the internal structure. Simple scoring systems, relying on number of errors and time taken, have proved to be reliable in other tests such as the FLS²². The rating can be made either while the tasks are being performed or afterwards from a video recording. Using video recordings allows blinded rating, which minimizes bias from raters.

Evidence of validity was found from relationships with other variables, as there were significant differences in performance scores between the groups with different levels of experience. Correlation was seen between the level of laparoscopic experience measured by the number of procedures performed and test performance scores. Measuring laparoscopic skills by the number of procedures can be problematic owing to recall bias. However, the number of procedures performed by a surgeon has been shown to correlate well with performance level and patient outcome². Having chosen 100 procedures as the cut-off criterion for experienced surgeons, it seemed reasonable to assume that all experienced surgeons had a sufficient level of competency regarding basic laparoscopic skills.

The pass–fail level was established using the contrasting groups method. Two of 20 novices passed the test, whereas two of 20 experienced surgeons failed. This pass–fail level seems acceptable as it discriminates well between competent and non-competent surgeons. As a result of this, some novices will pass the test after a short training period, and some experienced surgeons will fail. The FLS is also subject to a similar challenge; 18 per cent of competent surgeons are expected to fail the test and 18 per cent of non-competent surgeons to pass¹⁷. When setting a pass–fail level, it is important that it should be achievable by novice trainees⁹. To examine this further, it will be necessary to do more research exploring performance curves of the TABLT test among novice trainees.

Evidence of validity was explored by using a contemporary framework of validity¹⁴. An essential part of this process was to consider threats to validity²³. Threats to validity can be divided into two main categories: construct under-representation and construct irrelevant variance. Construct under-representation occurs when the content of the test does not sufficiently represent the construct. Construct irrelevant variance is a result of systematic bias. An example of a threat from construct under-representation is low reliability. In the present study,

a high ICC was found, indicating that the scoring system of the TABLT test was reliable.

Threats to validity from construct irrelevant variance include: rater-related bias, level of difficulty of the test and unjustifiable methods for setting the pass–fail level²³. The TABLT test can be rated by both direct observation and video recordings. A previous study²⁴ showed that experienced physicians were more highly rated as the result of direct observations than they were for blinded rating. The opposite was true for novices. This type of bias did not seem to influence the TABLT test scores, as demonstrated by the high ICC value. The reason lies in the simplicity of the scoring system, which used only time taken and number of errors made. The scoring system was easy to apply and therefore reduced the risk of rater-related bias. The variety of performance scores by novices and intermediates (*Fig. 3*) demonstrated that the level of difficulty of the test was appropriate. The present study also included a wide variety of experienced surgeons with different levels of experience. Using participants with many skill levels for the test is important, as wide variation in skill levels has been recognized among practising laparoscopic surgeons⁹.

The study included a large number of participants with a wide range of competencies across three surgical specialties. This is a strength of the study because all participants were either practising laparoscopic surgeons or training to be laparoscopic surgeons. The main limitation was the focus on the assessment aspect of the TABLT test. The training element of TABLT still needs to be investigated in more detail. In particular, the standard setting should be explored by allowing trainees to train on TABLT and examine their performance curves. The pass–fail level might be set too low, resulting in trainees reaching an insufficient level of competency. If the level were to be set too high, however, the test would not be fair, and trainees would spend time overtraining, which may not result in a higher level of skill in the operating theatre. This highlights a further limitation, as the transfer of skills from the TABLT to a clinical setting has not been explored. The scoring system, although easy to use and reliable, may be too simple to provide meaningful feedback to the trainees on their performance. Using a preformatted spreadsheet makes score calculations easy and transparent so that trainees themselves may be able to perform self-assessment during training or receive formative feedback from faculty. To ensure that appropriate laparoscopic techniques are acquired, the TABLT scores could be supplemented by feedback or ideas for improvement from an experienced surgeon.

Training models based on FLS have been developed for each surgical specialty^{19,25–27}. Although developed for general surgeons, research has demonstrated the benefits of

the FLS practical test in both urology and gynaecology^{28,29}. The TABLT test was designed to be used in cross-specialty training. The techniques practised are used widely, making it possible for experienced laparoscopic surgeons from one specialty to supervise a trainee from a different specialty. A cross-specialty approach also allows trainees from different specialties to practise together and learn from one another. An added benefit is that hospital departments and simulation centres can pool resources, offering more frequent courses and additional training opportunities. When trainees have mastered basic skills in laparoscopy, they can move on to specialty-specific training, such as operation modules on a virtual reality simulator⁴.

TABLT was developed to be used on a portable tablet trainer¹⁸ (Fig. 2), but may be used in any box trainer. Training on a portable trainer increases the flexibility of training³⁰. Box trainers are in general cost-effective, and have proved to be equally as good as virtual reality trainers for training basic laparoscopic skills³¹. Easy access to basic testing and training in laparoscopy could benefit trainees and, more importantly, their patients.

Acknowledgements

Funding and equipment was provided by the Centre for Clinical Education, Capital Region of Denmark. The instruments used in this trial were donated to the Centre for Clinical Education by Johnson & Johnson. E.T. has received non-restricted educational grants from Johnson & Johnson outside the submitted work.

Disclosure: The authors declare no other conflict of interest.

References

- 1 Darzi SA, Munz Y. The impact of minimally invasive surgical techniques. *Annu Rev Med* 2004; **55**: 223–237.
- 2 Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR *et al.*; Michigan Bariatric Surgery Collaborative. Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013; **369**: 1434–1442.
- 3 Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg* 1998; **175**: 482–487.
- 4 Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C *et al.* Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ* 2009; **338**: b1802.
- 5 Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004; **91**: 146–150.
- 6 Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA III *et al.* Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007; **193**: 797–804.
- 7 Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK *et al.* Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002; **236**: 458–463.
- 8 Reznick RK, MacRae H. Teaching surgical skills – changes in the wind. *N Engl J Med* 2006; **355**: 2664–2669.
- 9 Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G *et al.* Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005; **241**: 364–372.
- 10 Epstein RM. Assessment in medical education. *N Engl J Med* 2007; **356**: 387–396.
- 11 Korndorffer JR Jr, Kasten SJ, Downing SM. A call for the utilization of consensus standards in the surgical education literature. *Am J Surg* 2010; **199**: 99–104.
- 12 Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ* 2003; **37**: 830–837.
- 13 Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 2006; **119**: 166: e7–e16.
- 14 American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). *Standards for Educational and Psychological Testing*. American Educational Research Association: Washington, 2014.
- 15 Soper NJ, Fried GM. The fundamentals of laparoscopic surgery: its time has come. *Bull Am Coll Surg* 2008; **93**: 30–32.
- 16 Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. *Acad Med* 2014; **89**: 387–392.
- 17 Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc* 2003; **17**: 964–967.
- 18 Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P. Tablet based simulation provides a new solution to accessing laparoscopic skills training. *J Surg Educ* 2013; **70**: 161–163.
- 19 Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *J Gastrointest Surg* 2008; **12**: 210–212.
- 20 Molinas CR, De Win G, Ritter O, Keckstein J, Miserez M, Campo R. Feasibility and construct validity of a novel laparoscopic skills testing and training model. *Gynecol Surg* 2008; **5**: 281–290.
- 21 Schreuder HW, van den Berg CB, Hazebroek EJ, Verheijen RH, Schijven MP. Laparoscopic skills training using inexpensive box trainers: which exercises to choose when

- constructing a validated training course. *B7OG* 2011; **118**: 1576–1584.
- 22 Vassiliou MC, Ghitulescu GA, Feldman LS, Stanbridge D, Leffondre K, Sigman HH *et al.* The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. *Surg Endosc* 2006; **20**: 744–747.
- 23 Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ* 2004; **38**: 327–333.
- 24 Konge L, Vilmann P, Clementsen P, Annema JT, Ringsted C. Reliable and valid assessment of competence in endoscopic ultrasonography and fine-needle aspiration for mediastinal staging of non-small cell lung cancer. *Endoscopy* 2012; **44**: 928–933.
- 25 Tjiam IM, Persoon MC, Hendriks AJ, Muijtjens AM, Witjes JA, Scherpbier AJ. Program for laparoscopic urologic skills: a newly developed and validated educational program. *Urology* 2012; **79**: 815–820.
- 26 Kolkman W, van de Put MAJ, Woterbeek R, Trimboos JBMZ, Jansen FW. Laparoscopic skills simulator: construct validity and establishment of performance standards for residency training. *Gynecol Surg* 2008; **5**: 109–114.
- 27 Hennessey IA, Hewett P. Construct, concurrent, and content validity of the eoSim laparoscopic simulator. *J Laparoendosc Adv Surg Tech A* 2013; **23**: 855–860.
- 28 Dauster B, Steinberg AP, Vassiliou MC, Bergman S, Stanbridge DD, Feldman LS *et al.* Validity of the MISTELS simulator for laparoscopy training in urology. *J Endourol* 2005; **19**: 541–545.
- 29 Zheng B, Hur HC, Johnson S, Swanström LL. Validity of using Fundamentals of Laparoscopic Surgery (FLS) program to assess laparoscopic competence for gynecologists. *Surg Endosc* 2010; **24**: 152–160.
- 30 Korndorffer JR Jr, Bellows CF, Tekian A, Harris IB, Downing SM. Effective home laparoscopic simulation training: a preliminary evaluation of an improved training paradigm. *Am J Surg* 2012; **203**: 1–7.
- 31 Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 2013; **257**: 586–593.

Ensuring Competency of Novice Laparoscopic Surgeons—Exploring Standard Setting Methods and their Consequences

Ebbe Thinggaard, MD,^{*,†} Flemming Bjerrum, MD,^{*} Jeanett Strandbygaard, PhD,[‡] Ismail Gögenur, DMSc,^{*} and Lars Konge, PhD[†]

^{*}Department of Surgery, Roskilde and Koege Hospital, Koege, Denmark; [†]Copenhagen Academy for Medical Education and Simulation, Rigshospitalet, Copenhagen, Denmark; and [‡]Department of Gynecology, Juliane Marie Centre, Rigshospitalet, Copenhagen, Denmark

OBJECTIVE: Simulation-based assessment tools have been developed to allow for proficiency-based simulator training in laparoscopy. However, few studies have examined the consequences of different standard setting methods or examined what level of proficiency is considered adequate for trainees. The objectives of the present study were to explore the consequences of different standard setting methods and to examine the proficiency level that surgical trainees are expected to reach, before performing supervised surgery on patients.

DESIGN: Study participants undertook the Training and Assessment of Basic Laparoscopic Techniques test. The tests were video-recorded and rated using a simple scoring system based on number of errors and time. Participants were then asked to assess how high a score a novice should reach before performing supervised surgery on a patient. We then compared 3 methods of standard setting: expert performance level, contrasting groups method, and a modified Angoff method.

SETTING: The study was conducted at the Copenhagen Academy for Medical Education and Simulation. The academy provides surgical simulation training in laparoscopy for trainees at the hospitals in the Capital Region and the Zealand Region of Denmark.

PARTICIPANTS: Participants were recruited among surgical trainees in their first year of specialty training from surgery, gynecology, and urology departments. A total of 40 participants were included and completed the trial.

RESULTS: The different standard setting methods resulted in different pass/fail levels. At the expert performance level,

the pass/fail level was 474 points—the contrasting groups method resulted in 358 points and the modified Angoff method resulted in 311 points among experienced surgeons, and 386 points among trainees. The different proficiency levels resulted in a failure rate of 0% to 50% of experienced surgeons and a pass rate of 0% to 25% of novices. Novice laparoscopic surgeons set a higher pass/fail level than experienced surgeons did ($p = 0.008$).

CONCLUSION: Required proficiency levels varies depending on the standard setting method used, which highlights the importance of using an established standard setting method to set the pass/fail level. (J Surg Ed 1:111-111. © 2016 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: laparoscopy, minimally invasive surgery, simulation, training, standard setting, medical education

COMPETENCIES: Practice-Based Learning and Improvement

INTRODUCTION

The traditional Halsteadian approach to training surgeons includes having novice trainees participate in surgery and operating on patients. Recent research has questioned this method, as trainee participation in operations can prolong operations and affect patient outcomes, with an increased risk of postoperative complications.¹ Simulation-based training has been suggested as a way of improving surgical training by creating “pretrained novices.” This approach has been shown to have a beneficial effect on patient outcomes, such as reducing the risk of intraoperative and postoperative complications and reducing the number of errors, operative time, and length of stay.²⁻⁶ There are currently

Correspondence: Inquiries to Ebbe Thinggaard, Department of Surgery, Roskilde and Koege Hospital, Lykkebackvej 1, 4600 Koege, Denmark; E-mail: ebbe.thinggaard@gmail.com, etg@regionsjaelland.dk

several simulation-based tests for laparoscopic surgery, such as Fundamentals of Laparoscopic Surgery⁷ for general surgeons, Laparoscopic Skills Testing and Training⁸ for gynecologists, Program for Laparoscopic Urological Skills⁹ for urologists, and Training and Assessment of Basic Laparoscopic Techniques (TABLT)¹⁰ developed for all surgical specialties. Proficiency-based simulation training has gained ground in laparoscopic surgical skills training as simulation-based assessment has developed.^{11,12} Setting a pass/fail level is a prerequisite for summative assessment of competency in proficiency-based training¹³ and, along with exploring the consequences, is an essential part of gathering validity evidence for a simulation-based test.^{14,15} However, defining the level of proficiency can be challenging and has been a focus of discussion in the literature of medical education and testing.^{13,16-18} The decision on a standard setting method is ultimately a policy decision as there is no true pass score for a test.¹³

There are a variety of standard setting methods that use criterion-based standard setting methods to assess competency.¹⁹ Criterion-based methods are either examinee-centered or test-centered.¹⁷ Examinee-centered methods determine the ability of the students and use these observations to set a pass/fail level. Test-centered methods look at the test characteristics, such as difficulty and relevance, and set a pass/fail level according to these characteristics. To explore the consequences of different standard setting methods, we chose to compare 3 standard setting methods. We compared the average expert performance levels, the contrasting groups method, and the modified Angoff method. The average expert performance level is a method in which the pass/fail level is set at the median performance level of a group of experienced surgeons.²⁰ The contrasting groups method¹⁷ sets the pass/fail level using the normal distribution of performance scores from 2 groups: competent and noncompetent. The pass/fail level is usually set at the intersection between the distributions of the 2 groups. The modified Angoff method consists of first asking judges to define a borderline student²¹ and then determine the performance level of a borderline student on each item in a test. The items scores are averaged across different judges and a pass/fail level is set.^{17,19} The Angoff method can include actual performance data from students presented to the judges, and it can be done through several iterations.²²

Laparoscopic surgical training has involved different standard setting methods, including receiver operator curves,⁷ a generalized examinee-centered method,²³ expert performance levels,²⁴ and the contrasting groups method.¹⁰ However, few studies looked at the consequences of the pass/fail setting or the effect of the choice of the standard setting method.²⁵ The present study aimed to explore the consequences of different standard setting methods and to examine what level of competency was

perceived to be adequate to begin performing supervised surgery.

MATERIAL AND METHODS

We used the TABLT test to explore our aim-of-study. The TABLT test is a training and testing tool developed for cross-specialty training in basic laparoscopic skills, including surgery, gynecology, and urology.¹⁰ It consists of 5 basic tasks, which can be practiced on a portable laparoscopic trainer.²⁶ The tasks include basic hand-eye coordination, cutting, sharp dissection, blunt dissection, and an integrated task simulating a cyst removal. Task scores are calculated by taking the maximum time of 600 seconds and then subtracting the time spent on each task and an error-specific penalty. Errors are defined for each task and include dropping a bead, cutting outside a circle, or perforating a balloon. Each error adds a 20-second penalty in the scoring system. Using scores from a group of expert laparoscopic surgeons, a standardized task score is calculated for each task. The performance score is the sum of the 5 standardized task scores and range from 0 points to 708 points.

Participants from departments of surgery, gynecology, and urology were recruited as part of the process of gathering evidence for the TABLT test.¹⁰ In total, 20 novices and 20 experienced laparoscopic surgeons participated in the study (Table). Novices were surgical trainees who had no prior operative experience in laparoscopy and less than 2 hours of experience practicing laparoscopy on a simulator. The experienced laparoscopic surgeons had performed more than 100 laparoscopic surgeries. All participants were asked to perform the TABLT test twice. The first attempt was done to help participants get used to the simulator and familiarize themselves with the TABLT tasks and the scoring system. The second attempt was video-recorded and rated using a simple scoring system based on number of errors and time. The ratings were recorded in a password-protected spreadsheet. The tests were supervised by one of the researchers (E.T.), who rated the tests on site. The videos were also rated by a blinded rater and used as part of establishing evidence of validity for the TABLT test.

TABLE. Participants Characteristics

	Novice	Experienced	Total
Number	20	20	40
Sex			
Male	7	13	20
Female	13	7	20
Range of age in years	24-31	31-58	24-58
Specialty			
Surgery	11	10	21
Urology	3	5	8
Gynecology	6	5	11

Based on the ratings, test scores were then used to calculate pass/fail levels using the expert performance levels and contrasting groups method. To establish a pass/fail level using the modified Angoff method, the experienced surgeons were asked how well a borderline surgical trainee should perform on the test. Specifically, they were asked how much time a borderline novice trainee could spend on each task, and the number of errors they would be allowed to make and still pass the test. Experienced surgeons were asked verbally after having finished the 2 attempts on the TABLT test; the answers were recorded by one of the researchers (E.T.). Novices were also asked how well a borderline surgical trainee should perform on the test, to examine the differences in the expected pass/fail level between experienced laparoscopic surgeons and novices.

Statistical analysis

To explore the consequences of each pass/fail level, the resulting pass/fail levels were evaluated by calculating the percentage of experienced surgeons who failed the test and the percentage of novices who passed the test. Differences in what level of competency was perceived to be adequate to begin performing supervised surgery were examined using an independent samples *t*-test. $p < 0.05$ was considered statistically significant. For analysis, statistical software was used (SPSS, version 20.0 Chicago, IL). The study was submitted for evaluation by the regional ethics committee, which determined that no approval was needed for the study (H-3-2013-FSP66).

RESULTS

All participants performed the TABLT test twice; the second attempt was video-recorded and rated. Each participant then assessed what level of proficiency they would expect a surgical trainee to reach before performing supervised surgery on a patient. An external blinded rater rated

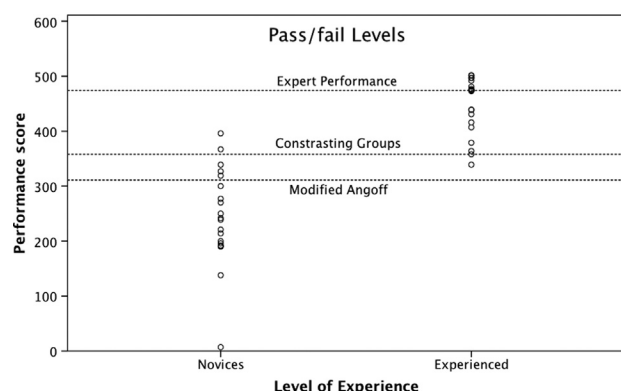


FIGURE 1. Pass/fail levels set by the different standard setting methods. Dotted lines indicate pass/fail levels using the 3 different standard setting methods.

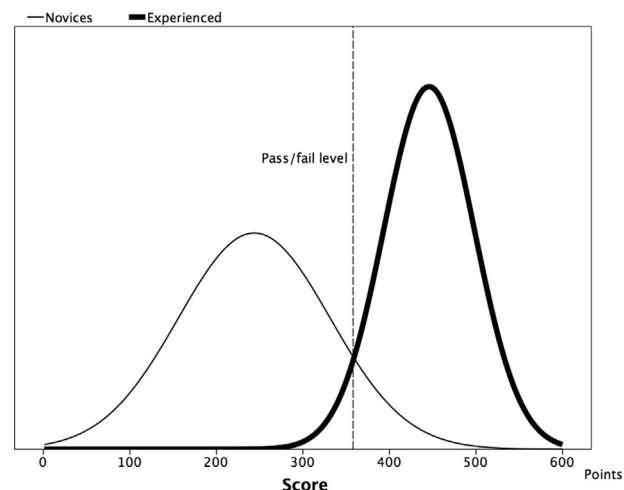


FIGURE 2. Pass/fail level set by contrasting groups method. The graphs indicate the distributions of scores from experienced laparoscopic surgeons and novice laparoscopic surgeons. The dotted line indicates the pass/fail level.

the 40 videos. Totally, 3 different methods of standard setting were used to establish pass/fail levels. The Table shows baseline characteristics of the participants.

The 3 standard setting methods used were expert performance level, contrasting groups method, and a modified Angoff method. Figure 1 shows the 3 resulting pass/fail levels. The expert performance level method was set at the median score of the experienced surgeons and resulted in the pass/fail level being set at 476 points (standard deviation [SD] 53). At this level, 50% of experienced surgeons would fail the test and no novices would pass the test. Using the contrasting groups method, a pass/fail level was set at the intersection between the normal distributions of experienced surgeons and novices. This

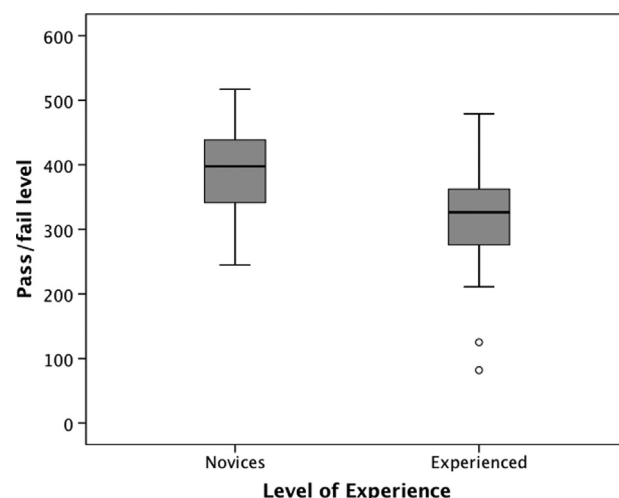


FIGURE 3. The expected pass/fail level set by novices and experienced surgeons indicating the level that they would expect a borderline novice laparoscopic surgeon to be able to reach. Boxes indicate IQR, lines indicate median values, and circles indicate outliers. IQR, interquartile range.

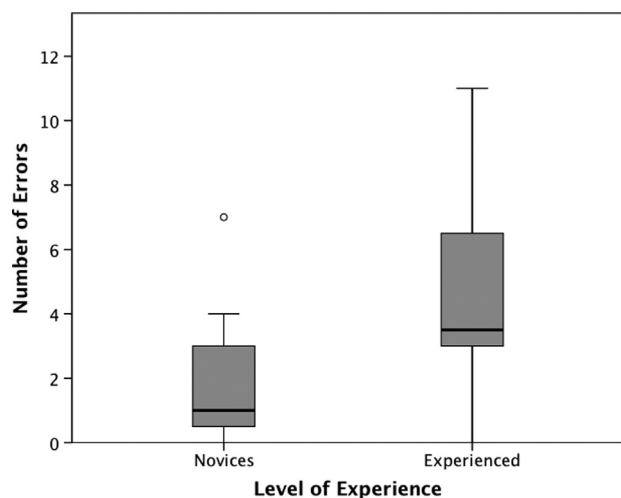


FIGURE 4. The number of errors acceptable to still pass the test, assessed by novices and experienced surgeons, respectively. The number of errors indicates the highest number of errors they would expect a borderline novice laparoscopic surgeon to make when passing the test. Boxes indicate IQR, lines indicate median values, and circles indicate outliers. IQR, interquartile range.

resulted in a pass/fail level set at 358 points. At this level, 10% of experienced surgeons would fail the TABLT test and 10% of novices would pass (Fig. 2). When using the Angoff method to decide on the pass/fail level, both experienced surgeons and novices were asked how high a score they would expect a borderline novice laparoscopic surgeon to have to reach to pass the test. Experienced surgeons set a pass/fail level at 311 points (SD 94) and novices set the pass/fail level at 386 points (SD 78) (Fig. 2). At 311 points, no experienced surgeons would fail and 25% of novices would pass the test. At 386 points, 20% of experienced surgeons would fail the test and 5% of novices would pass the test. When deciding what the pass/fail level should be, both experienced surgeons and novices were asked to assess the pass/fail level using the scoring system based on number of errors and time. There was a significant difference between experienced and novices in what level of competency was perceived to be adequate to begin performing supervised surgery ($p = 0.008$). The number of errors was the deciding factor; novices expected a pass/fail level with a mean number of errors of 1 (range: 0-7). Experienced surgeons expected a pass/fail level with a mean number of errors of 3.5 (range: 0-11); $p = 0.001$ (Figs. 3 and 4).

DISCUSSION

It can be challenging to set a pass/fail level for an assessment tool because different methods of standard setting can be used. In the present study, we compared 3 different types of standard setting methods and found 3 different pass/fail levels that resulted in different passing and fail rates among experienced and novice laparoscopic surgeons. The

consequences of different pass/fail setting varied between failing 0% to 50% of experienced surgeons and passing 0% to 25% of novices. We also found that there was a significant difference between experienced and novices in what level of competency was perceived to be adequate to begin performing supervised surgery.

We recruited a variety of laparoscopic surgeons from 3 different specialties, with different levels of skills, which increased the generalizability of the study. Having laparoscopic surgeons with varying levels of skills reflected the real variety of skills among practicing laparoscopic surgeons and is also a vital part of establishing a credible pass/fail level.¹² The limitations of our study include the fact that the sample size was small, which reduced the external validity of our findings. Participants were only asked to perform 2 attempts on the TABLT test. In total, 2 attempts may have been insufficient for some laparoscopic surgeons who were not used to simulator or the tasks to demonstrate their actual level of proficiency. Allowing participants more attempts could have been beneficial, as performance curves could have been observed and the variation in performance by the experienced groups may have been reduced. This would also have allowed an exploration of the training aspect of the TABLT test. Performance curves would have also given greater insight into the fairness of the pass/fail level, as we would have been able to see whether novices were able to reach the pass/fail level after training. To explore the fairness of the pass/fail level, the study could also have included a test of transference, which could be examined using objective rating scales of a laparoscopic procedure, or as a surrogate, using a procedure-specific module on a virtual reality simulator.²⁷

The appropriate proficiency level of trainees before performing supervised surgery is unknown. In the present study, we looked at what novices and experienced laparoscopic surgeons expected of trainees. Experienced surgeons set a lower pass/fail than novices did and a lower level than any of the other standard setting methods. This study has shown that the number of errors deemed acceptable was the deciding factor. Experienced laparoscopic surgeons were more accepting of trainees making errors than novices were. This is a surprising result, as experts tend to set pass/fail levels quite high.¹³ Previous studies in medical education have also found differences when comparing different standard setting methods.²² The question remains regarding which standard setting to use and what pass/fail level is fair. When choosing a standard setting method, it is important that the method is well supported in previously published literature, and that the method provides a pass/fail level that is credible to trainees taking the test and to the faculty administering the test.¹⁸ The pass/fail level should not be set too high or too low.¹³ A high pass/fail level results in unfair testing; in order for the test to be fair, the proficiency level must be reachable for trainees when acquiring laparoscopic skills.¹² A low pass/fail level could result in low levels of skills being acquired and thereby missing out on the training potential in simulation training. A low pass/fail

level could diminish the effect of simulation-based training, as trainees would have to focus on learning basic skills during a supervised operation. During proctored clinical training, they should be able to learn more complex skills, such as decision-making, communication, and other non-technical skills, which are vital in the operating theater.

Standard setting remains a policy decision, as there is no true pass/fail score for a test.^{13,28} However, standard setting methods should be judged on how well they work,¹³ and a pass/fail level should demonstrate the relationship between test scores and performance in practice.¹⁸ Relying on test performance by experienced surgeons as a source of validity evidence may not be sufficient for assessing a standard setting method or setting the pass/fail level. An important part of contemporary approaches to validity is the consequences of testing.¹⁴ Looking at actual performance in the clinical setting would provide better information about the relevance of the standard setting method and the pass/fail level. Many studies have looked at the construct validity of an assessment tool and the benefits derived from training on different types of simulators.²⁹ However, few studies have examined the effects of different pass/fail levels, and the consequences of proficiency levels of assessment tools, on actual operative performance. Research into the transfer of skills depending on different standard setting and pass/fail levels could provide valuable information about the level of competency we should expect trainees to reach when training on simulators.

CONCLUSIONS

Standard setting is a vital part of competency-based training as it defines the level of proficiency that is required to be reached. We found that different standard setting methods resulted in different pass/fail levels. The level must be fair to trainees, and therefore achievable, but also high enough to ensure that the necessary level of competency is reached. In the present study, experienced surgeons set a lower pass/fail level than novices did, revealing a difference in how the role of simulation training is perceived by novices and experienced laparoscopic surgeons.

ACKNOWLEDGMENTS

The study was supported by a nonrestricted educational grant from Johnson & Johnson, United States. Johnson & Johnson had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

REFERENCES

1. Kasotakis G, Lakha A, Sarkar B, et al. Trainee participation is associated with adverse outcomes in emergency general surgery: an analysis of the National Surgical Quality Improvement Program database. *Ann Surg.* 2014;260(3):483-490 discussion 490-483.
2. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg.* 2002;236(4):458-463 discussion 463-454.
3. Zendejas B, Cook DA, Hernandez-Irizarry R, Huebner M, Farley DR. Mastery learning simulation-based curriculum for laparoscopic TEP inguinal hernia repair. *J Surg Educ.* 2012;69(2):208-214.
4. Larsen CR, Soerensen JL, Grantcharov TP, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *Br Med J.* 2009; 338(b1802):1253.
5. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg.* 2004;91(2):146-150.
6. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg.* 2007;193(6):797-804.
7. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc.* 2003;17(6):964-967.
8. Molinas CR, De Win G, Ritter O, Keckstein J, Miserez M, Campo R. Feasibility and construct validity of a novel laparoscopic skills testing and training model. *Gynecol Surg.* 2008;5(4):281-290.
9. Tjiam IM, Persoon MC, Hendrikx AJ, Muijtjens AM, Witjes JA, Scherpbier AJ. Program for laparoscopic urologic skills: a newly developed and validated educational program. *Urology.* 2012;79(4):815-820.
10. Thinggaard E, Bjerrum F, Strandbygaard J, Gogenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *Br J Surg.* 2015; 102(9):1106-1113.
11. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med.* 2006;355(25): 2664-2669.
12. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg.* 2005;241(2):364-372.
13. Kane MT, Crooks TJ, Cohen AS. Designing and evaluating standard-setting procedures for licensure and certification tests. *Adv Health Sci Educ Theory Pract.* 1999;4(3):195-207.

14. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ*. 2003;37(9):830-837.
15. Ghaderi I, Manji F, Park YS, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. *Ann Surg*. 2015;261(2):251-262.
16. McKinley DW, Norcini JJ. How to set standards on performance-based examinations: AMEE Guide No. 85. *Med Teach*. 2014;36(2):97-110.
17. Ben-David MF. AMEE Guide No. 18: standard setting in student assessment. *Med Teach*. 2000;22(2):120-130.
18. Norcini JJ, Shea JA. The credibility and comparability of standards. *Appl Meas Educ*. 1997;10(1):39-59.
19. Norcini JJ. Setting standards on educational tests. *Med Educ*. 2003;37(5):464-469.
20. Yudkowsky R, Park YS, Lineberry M, Knox A, Ritter EM. Setting mastery learning standards. *Acad Med*. 2015;90(11):1495-1500.
21. Angoff WH. Scales, norms and equivalent scores. Thorndike RL, editor. *Educational Measurement*. 2nd ed. Washington: American Council on Education, 1971. p. 508-600.
22. Cusimano MD. Standard setting in medical education. *Acad Med*. 1996;71(10):S112-S120.
23. Tjiam IM, Schout BM, Hendrikx AJ, et al. Program for laparoscopic urological skills assessment: setting certification standards for residents. *Minim Invasive Ther Allied Technol*. 2013;22(1):26-32.
24. Kolkman W, van de Put MAJ, Woterbeek R, Trimbos JBMZ, Jansen FW. Laparoscopic skills simulator: construct validity and establishment of performance standards for residency training. *Gynecol Surg*. 2008;5(2):109-114.
25. Zendejas B, Ruparel RK, Cook DA. Validity evidence for the Fundamentals of Laparoscopic Surgery (FLS) program as an assessment tool: a systematic review. *Surg Endosc*. 2015;20:1-9.
26. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P. Tablet based simulation provides a new solution to accessing laparoscopic skills training. *J Surg Educ*. 2013;70(1):161-163.
27. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking simulation-based educational assessments and patient-related outcomes: a systematic review and meta-analysis. *Acad Med*. 2015;90(2):246-256.
28. Jaeger RM. Certification of Student Competence. Linn RL, editor. *Educational Measurement*. 3rd ed. New York: American Council on Education and Macmillan, 1989. p. 485-514.
29. Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg*. 2013;257(4):586-593.

TITLE PAGE:

Title: Take-Home Training in Laparoscopy Facilitates Distributed Training, a Randomized Trial

Authors:

1: Ebbe Thinggaard (ET), MD, Copenhagen Academy for Medical Education and Simulation, Capital Region, Denmark, Copenhagen, Denmark and Department of Surgery, Zealand University Hospital, Koege, Denmark.

2: Flemming Bjerrum (FB), MD, Department of Surgery, Zealand University Hospital, Koege, Denmark.

3: Jeanett Strandbygaard (JS), PhD, Department of Obstetrics and Gynecology, Rigshospitalet, Copenhagen, Denmark.

4: Lars Konge (LK), PhD, Copenhagen Academy for Medical Education and Simulation, Capital Region, Denmark, Copenhagen, Denmark

5: Ismail Gögenur (IG), DMSc, Department of Surgery, Zealand University Hospital, Denmark.

Author responsible for correspondence and reprints:

Ebbe Thinggaard, Department of Surgery, Zealand University Hospital, Lykkebaekvej 1, 4600 Koege, Denmark, Tel: +45-52722546, ebbe.thinggaard@gmail.com

Source of Funding:

Equipment was provided by Copenhagen Academy for Medical Education and Simulation, Capital Region, Denmark, Copenhagen, Denmark

Dr. Ebbe Thinggaard and Ismail Gögenur report non-restricted educational grants from Johnson and Johnson during the conduct of the submitted work.

Drs. Flemming Bjerrum, Jeanett Strandbygaard and Lars Konge have no conflicts of interest or financial ties to disclose.

Category: Original Communications.

ABSTRACT:

Background:

Simulation training in laparoscopy has become a valuable addition to supervised clinical training. However, barriers to simulation training still exist. Training at home on mobile boxtrainers make training more accessible and enables participants to organize training according to their own schedule. The objectives of this study were to examine the added effects of training at home during a laparoscopic training program at a simulation center and to explore participant's ability to rate their own performance.

Methods:

Participants were recruited among trainees taking part in a laparoscopic training course. The intervention consisted of a mobile boxtrainer for training at home. All participants also had access to training on virtual reality simulators and were given a logbook to keep record of their training.

Results:

A total of 36 participants completed the trial. There was a significant difference in the number of training sessions used (5.8 vs. 2.3, $p < 0.001$), but no difference in the number of days (86 days vs. 89 days, $p = 0.89$), or time spent training on boxtrainers (302 minutes vs. 218 minutes, $p = 0.26$). There was no difference on the final test score (493 vs. 460, $p = 0.07$). Participants were able to rate their own performance reliably, ICC 0.86, $p < 0.001$.

Conclusions:

Training at home on mobile boxtrainers resulted in a more distributed training pattern where participants trained more frequent and in shorter sessions. Participants were able to rate their own performance during unsupervised training at home.

INTRODUCTION

The approach to surgical training has changed dramatically. In many places training on simulators has become part of how surgical trainees are prepared for surgery. Training in a simulated environment has shown to improve patient outcomes¹⁻⁴ and is a valuable addition to the traditional Halsteadian method of training surgeons at the operating table. Simulation training is now a requirement to obtain specialty registration in the USA⁵ and much of the initial training of basic laparoscopic skills can be done in simulation centers. Although many surgical trainees and their patients have benefitted from these developments, barriers to simulation training still exist. Studies exploring the use of simulation training in laparoscopic surgery have identified barriers such as access to simulators, time for training, and financial constraints—as simulation equipment can be costly.⁶ Furthermore, training at simulation centers is often organized as massed training sessions in boot camps where there is no follow-up on training. To overcome the barriers in simulation training, simple mobile boxtrainers (BT) have been developed which allow training at home on affordable simulators at a time that suits the trainee.⁷ Nonetheless, training at home without supervision poses new challenges, as feedback from faculty is not readily available.⁸ Home training of laparoscopic skills has shown to be feasible and to affect training by introducing a more individualized approach where trainees vary their training by shifting between different tasks during training.⁹ However, providing trainees with the freedom to organize their training could also change training patterns, allowing for more distributed training where trainees practice more frequently in shorter intervals. An approach which is beneficial for technical skills acquisition^{10, 11} and in line with educational principles of deliberate practice¹² and Directed Self Regulated Learning (DSRL).¹³ Many simulation centers nowadays offer training programs using either virtual reality simulators (VRS) or boxtrainers (BT) for in-center simulation training. Home training could be a valuable addition to current training programs and there is a need to investigate how training at home can be implemented and improve current training programs. We designed a study to meet the need for guidance on implementation of home training in laparoscopic skills training. The purpose of this study was to examine the effects of training at home on course participants training patterns. We looked at the numbers of days it took to complete the training, time spent on training, and

number of training sessions. We also looked at whether training at home on a boxtrainer would decrease the time spent training on a VRS in the simulation center. A secondary aim was to explore participants' ability to rate their own performance when training unsupervised, using a structured self-rating system.

METHODS

Setting:

At the Copenhagen Academy for Medical Education and Simulation (CAMES),¹⁴ doctors in specialty training participate in a basic laparoscopic skills training program during the first year of their training. The course is a cross-specialty training program for doctors from the departments of gynecology, urology and surgery.¹⁵ The aim of the course is to prepare course participants for their first supervised laparoscopic surgical procedure. The course consists of two formalized one-day courses separated by a period of self-directed training on VRS and BT (Figure 1). The first part of the program is an introductory course which includes theoretical teaching in traditional classroom style mixed with practical sessions to prepare trainees for training on VRS and BT. After the introduction course the participants go through a period of self-directed training where they book training sessions at the simulation center, and practice on both VRS and BT. At the simulation center they are assisted by a simulator technician who is able to give technical assistance and provide some feedback during training. As part of the self-directed training, participants are required to pass the Training and Assessment of Basic Laparoscopic Techniques¹⁶ test on the BT and to reach a predefined level of proficiency on the VRS. The TABLT test is a training and assessment system consisting of five simple tasks: peg-transfer, cutting, sharp dissection, blunt dissection and a cyst removal. Each task has specified type of errors, and a pass/fail level has been set so that the goal is clear for the trainees. When training on the TABLT participants can rate their own performance and see when they have reached the pass/fail level. To pass the TABLT test, participants first have to perform a pre-test where they rate their own performance using a simple scoring system based on time and number of errors¹⁶. When they have handed in their self-rated test, they then can book a time for a proctored test where a member of faculty is present during testing. After having completed the two compulsory elements of the course, reaching proficiency on the VRS and passing the TABLT test, participants are able to participate on the final operative course. Participants decide themselves when to sign up for the final course and are therefore free to organize and distribute their training at a schedule that suits their needs.

Participants:

The course participants consisted of doctors in the first year of their specialty training. Due to practical difficulties in organizing training to fit the individual need of participants, participants often have different levels of experience as some have started performing supervised laparoscopic surgeries while others have only assisted. Participants were invited to take part in the trial during a one-year period.

Intervention:

The intervention consisted of the addition of home training on a mobile BT. The intervention group trained at the simulation center but were also given a portable BT¹⁷ to practice at home. The control group only trained at the simulation center. Both groups had access to training on VRS at the simulation center.

Randomization:

The primary investigator (ET) was responsible for inclusion of participants. After enrolment, participants were randomly allocated using a computer generated allocation sequence (www.randomizer.org). The administrator at the simulation center retrieved the allocation sequence and kept the sequence concealed until the allocation had been finalized.

Outcomes:

All participants were given a training log to record their training. Based on information from the logbooks we looked at the number of days from enrolment to passing the TABLT test, the time spent training and the number of training sessions. Three outcomes were used; number of days, time spent training, and number of training sessions. The three outcomes were chosen to allow for an adequate analysis of changes in training patterns. The number of days was chosen as the primary outcome, and used for the sample size calculation, as it was anticipated that training at home would allow trainees to organize their training more effectively and result in participants being able to pass the TABLT in shorter period of time. We also explored differences in the performance levels that participants reached on the final TABLT test. To investigate the effect of training at home on VRS training taking

place at the simulation center we looked at the number of days from enrolment to reaching a predefined level on the VRS, time spent training on the VRS and the number of training sessions. To explore the participant's ability to rate them selves we looked at the reliability of participants rating their pre-test and the rating of a trained blinded rater.

Statistical analysis:

The sample size for the trial was calculated based on the assumption that the control group would pass the TABLT test after six weeks of practice (42 days), SD three weeks (21 days). The intervention group was expected to pass after four weeks of practice (28 days), SD of three weeks (21 days). Setting alpha at 0.05 and beta at 0.10, 24 participants were required in each group. Accounting for inaccuracies we expected to include a total of 50 participants in the trial. A recruitment period of one year was decided on, as this would allow for a decision on implementation of home training for the following year. During the year, six courses were planned and 74 doctors would be able to participate making the recruitment of 50 participants feasible. In case we did not manage to include sufficient participants, a plan was made to evaluate whether the trial should continue enrolling participants after one year. The plan consisted of making a new sample size calculation based on the available data to decide whether continuing inclusion would be beneficial. To analyze whether there was a significant level of difference in the above-mentioned measurements, student t-test would be used. A p-value below 0.05 was considered statistically significant for the primary outcome. As we included three outcomes, the primary and the two secondary outcomes would be analyzed and adjusted using the Bonferroni method. The Intraclass Correlation Coefficient was used to examine reliability of participants self-rating. A statistical software package was used (SPSS vs. 20.0, Chicago, IL).

Trial registration:

The trial was submitted for evaluation to the regional ethics committee, which determined that no approval was needed for the trial (H-3-2014-FSP31), the trial was also registered at www.clinicaltrials.gov prior to commencing the trial (NCT02243215), and the trial was performed according to the CONSORT statement.

RESULTS

We included participants during a one-year period in which 54 doctors participated in the training course. 50 participants were invited to participate of which 46 were enrolled in the study and 36 completed the training course during the study. Four participants dropped out of the training course and six participants were excluded from the study, as they did not complete the training course during the one-year study period. Out of the 36 who completed the course 18 were from to the control group, and 18 were from to the intervention group. For participants baseline characteristics see Table 1. We performed a new sample size calculation after the one-year study period, as we did not reach our goal of having 50 participants complete the training course. The calculation was based on data available from the 36 participants, corresponding to 75% of the anticipated sample size. Setting alpha at 0.05 and beta at 0.10, 11,422 participants in each group would be required, which was not deemed feasible and therefore we decided not to continue recruiting participants in the next one-year period.

We found a significant difference in the number of training sessions but no difference in the number of days from enrolment to passing the TABLT test, or on the time spent training, see table 2. We found no difference on the final TABLT test score when exploring differences in the performance level participants reached, (Table 2). Also there was no difference in the number of days from enrolment to reaching proficiency on the VRS, time spent training, or number of training session when investigating the effect of training at home on VRS training at the simulation center (Table 3). There was a good reliability when comparing participants' ratings of their pre-test and that of a blinded rater, ICC 0.86, $p < 0.001$. There was no difference in the passing rate between the two groups, all of the participants were able to reach the pass level on the TABLT test on the pre-test, and all but one participant passed the proctored test on the first try. One person failed their first attempt but passed on the second try.

DISCUSSION

We found a significant difference in the number of training sessions. Therefore our trial shows that participants training at home practice more frequently and in shorter intervals, as there was no difference in the number of days or time spent training to pass the TABLT test. Participants could reliably rate their own performance and 100% were able to pass the TABLT test on a pre-test using a structured self-rating system.

We investigated the effect of both training at home on a simple mobile BT and training at a simulation center. It was not possible to compare the effect of only using home-based training with that of training at a simulation center due to the study design. All participants had access to support from simulation technicians at the simulation center even the intervention group. Accordingly, implementing home-based training may be more challenging than our observations show. In the training program we use both VRS and BT; mixing two training methods could cloud findings. A trial focusing on BT exclusively could have more clearly demonstrated the benefits of training at home using a BT. However, examining the use of training at home as a supplement was a deliberate choice in the study design. We chose to do the study under realistic circumstances as part of an existing laparoscopic training program. Choosing this design generated findings that provide information on the effects of integrating home training in existing training programs. In the basic laparoscopy course we also use a cross-specialty approach to laparoscopic training where doctors from different specialties practice together, which increases the external validity as findings can be generalized across training programs for different specialties. Participants in our study had different levels of experience prior to commencing the training program. This makes the results of the trial applicable to trainees with different degrees of experience.

Having participants from different specialties and with different levels of experience can have impacted the results. When performing the initial power calculations we anticipated that six weeks of training was sufficient for the control group to pass the TABLT test and four weeks for the intervention group. Nonetheless, during the trial we found that the duration of training in general was much longer and that training patterns varied greatly among participants. These findings demonstrate that factors other than access to training are important determinants of training duration and training

patters. In our study, the final part of the training program was the operative course, which was on fixed dates six times a year. Participants decided themselves when to enroll for the final course but did so before reaching proficiency on the VRS and passing the TABLT test. This may have imposed a structure on training duration and patterns that influenced the self-directed part of the training course as the final course provided a deadline for when to pass the TABLT test. Accordingly, participants entered a training program governed by the date of the final operative course. Having a deadline could have a substantial effect on when trainees decided to practice, and on the duration of their training. A further analysis of training patterns and the effect of deadlines and obligatory tests would be needed to explore this in more depth. Nevertheless, the structure imposed by a deadline was the same for both groups, and in our trial we found that home training led to a more distributed training pattern where participants trained more frequent but in shorter training sessions. Distribution of training in shorter and more frequent training sessions has been shown to improve training outcomes compared with massed training sessions.^{10, 11} Distributed training is recommended for laparoscopic virtual reality simulator training¹⁸ and learning curves in particular has been shown to be improved by using distributed training when compared with massed training.¹⁹ Although the ideal training interval for laparoscopic simulation training has not been identified, it has been shown that short training intervals are superior to long training intervals.²⁰ Our study demonstrates that training at home gives trainees the freedom to practice more frequently and for shorter intervals, by allowing them to practice at their convenience. This importance of this finding is supported by a study showing that shorter but more frequent training sessions are beneficial in training of laparoscopic skills.²¹ The problem in massed training is fatigue. Fatigue influence psychomotor learning and have shown to impact the level of proficiency in skills training.²² Furthermore distributed training is also an important element of deliberate practice,¹² which one study has explored in home-based simulation training. The study demonstrated that training at home led to a more individualized approach where participants varied their training by changing between different tasks.⁹

An individualized approach to training has emerged as an instructional method called Directed Self Regulated Learning (DSRL),^{13, 23} which is recommended for simulation training.²⁴ Principles of DSRL have shown to be useful in both advanced cardiac life support skills training²⁵ and in VRS

mastoidectomy training.²⁶ For training of laparoscopic skills at home this approach could be of great value. When considering unsupervised laparoscopic skills training at home, using DSRL as a strategy would allow for a structured training program where trainees are in control of their training. This approach is well suited to help overcome the challenges in simulation training at home where supervision and feedback from faculty are unavailable. Future research should focus on implementing home training founded on DSRL theories as this would allow trainees to overcome barriers in simulation training, such as time and location, by providing a structured but self-regulated training program accessible on affordable mobile boxtrainers.

Conclusions:

Home-based training of basic laparoscopic skills on a mobile boxtrainer allowed trainees to practice at their own convenience and resulted in shorter and more frequent training sessions. Trainees could reliably rate themselves when using a structured self-rating system, which resulted in a 100% pass rate on a self-rated test.

REFERENCES

1. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004;**91**: 146-150.
2. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, Schroeder TV, Ottesen BS. Effect of virtual reality training on laparoscopic surgery: Randomised controlled trial. *BMJ* 2009;**338**: 1253.
3. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002;**236**: 458-463.
4. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, 3rd, Ramel S, Smith CD, Arvidsson D. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007;**193**: 797-804.
5. Soper NJ, Fried GM. The fundamentals of laparoscopic surgery: its time has come. *Bull Am Coll Surg* 2008;**93**: 30-32.
6. Caban AM, Guido C, Silver M, Rossidis G, Sarosi G, Ben-David K. Use of collapsible box trainer as a module for Resident Education. *J Soc Laparoendo* 2013;**17**: 440-444.
7. Partridge R, Hughes M, Brennan P, Hennessey I. There is a worldwide shortfall of simulation platforms for minimally invasive surgery. In: *Journal of Surgical Simulation*; 2015. p. 12-17.
8. van Empel PJ, Verdam MG, Strypet M, van Rijssen LB, Huirne JA, Scheele F, Bonjer HJ, Meijerink WJ. Voluntary autonomous simulator based training in minimally invasive surgery, residents' compliance and reflection. *J Surg Educ* 2012;**69**: 564-570.
9. Korndorffer Jr JR, Bellows CF, Tekian A, Harris IB, Downing SM. Effective home laparoscopic simulation training: A preliminary evaluation of an improved training paradigm. *Am J Surg* 2012;**203**: 1-7.
10. Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychol Sci Public Interest* 2013;**14**: 4-58.
11. Benjamin AS, Tullis J. What makes distributed practice effective? *Cognitive Psychology* 2010;**61**: 228-247.
12. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;**79**: S70-81.
13. Brydges R, Dubrowski A, Regehr G. A new concept of unsupervised learning: directed self-guided learning in the health professions. *Acad Med* 2010;**85**: S49-55.
14. Konge L, Ringsted C, Bjerrum F, Tolsgaard MG, Bitsch M, Sorensen JL, Schroeder TV. The Simulation Centre at Rigshospitalet, Copenhagen, Denmark. *J Surg Educ* 2015;**72**: 362-365.
15. Bjerrum F, Sorensen JL, Thinggaard E, Strandbygaard J, Konge L. Implementation of a Cross-specialty Training Program in Basic Laparoscopy. *JSLs* 2015;**19**.
16. Thinggaard E, Bjerrum F, Strandbygaard J, Gogenur I, Konge L. Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *Br J Surg* 2015;**102**: 1106-1113.
17. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P. Tablet based simulation provides a new solution to accessing laparoscopic skills training. *J Surg Educ* 2013;**70**: 161-163.
18. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;**241**: 364-372.

19. Andersen SA, Konge L, Caye-Thomasen P, Sorensen MS. Learning Curves of Virtual Mastoidectomy in Distributed and Massed Practice. *JAMA Otolaryngol Head Neck Surg* 2015; 1-6.
20. Stefanidis D, Walters KC, Mostafavi A, Heniford BT. What is the ideal interval between training sessions during proficiency-based laparoscopic simulator training? *Am J Surg* 2009;**197**: 126-129.
21. Gallagher AG, Jordan-Black JA, O'Sullivan GC. Prospective, randomized assessment of the acquisition, maintenance, and loss of laparoscopic skills. *Ann Surg* 2012;**256**: 387-393.
22. Kahol K, Leyba MJ, Deka M, Deka V, Mayes S, Smith M, Ferrara JJ, Panchanathan S. Effect of fatigue on psychomotor and cognitive skills. *Am J Surg* 2008;**195**: 195-204.
23. Brydges R, Nair P, Ma I, Shanks D, Hatala R. Directed self-regulated learning versus instructor-regulated learning in simulation training. *Med Educ* 2012;**46**: 648-656.
24. Brydges R, Manzone J, Shanks D, Hatala R, Hamstra SJ, Zendejas B, Cook DA. Self-regulated learning in simulation-based training: A systematic review and meta-analysis. *Medical Education* 2015;**49**: 368-378.
25. Devine LA, Donkers J, Brydges R, Perelman V, Cavalcanti RB, Issenberg SB. An Equivalence Trial Comparing Instructor-Regulated With Directed Self-Regulated Mastery Learning of Advanced Cardiac Life Support Skills. *Simul Healthc* 2015;**10**: 202-209.
26. Andersen SA, Foghsgaard S, Konge L, Caye-Thomasen P, Sorensen MS. The effect of self-directed virtual reality simulation on dissection training performance in mastoidectomy. *Laryngoscope* 2015.

Table 1: Participants Baseline Characteristics:

	Intervention group	Control group	Total
Number of participants	18	18	36
Age (median, range)	30 (25-36)	30 (25-46)	30 (25-46)
Gender			
Men / Women	5 / 13	5 / 13	10 / 26
Specialty			
Surgery	6	5	11
Urology	3	3	6
Gynecology	9	10	19
Dominant hand			
Right / Left	16 / 2	16 / 2	32 / 4

Table 2: Training on Boxtrainers

	Intervention group mean (95%CI)	Control group mean (95%CI)	p-value *significant
Number of days	86 (52-120)	89 (52-127)	p = 0.89
Time spent training, in minutes	302 (189-414)	218 (112-223)	p = 0.74
Number of training sessions	5.8 (4-7.5)	2.3 (1.5-3.1)	p < 0.001*
Final TABLT test score	493 (465-522)	460 (434-485)	p = 0.63

TABLT: Training and Assessment of Basic Laparoscopic Techniques.

Table 3: Training on Virtual Reality Simulator

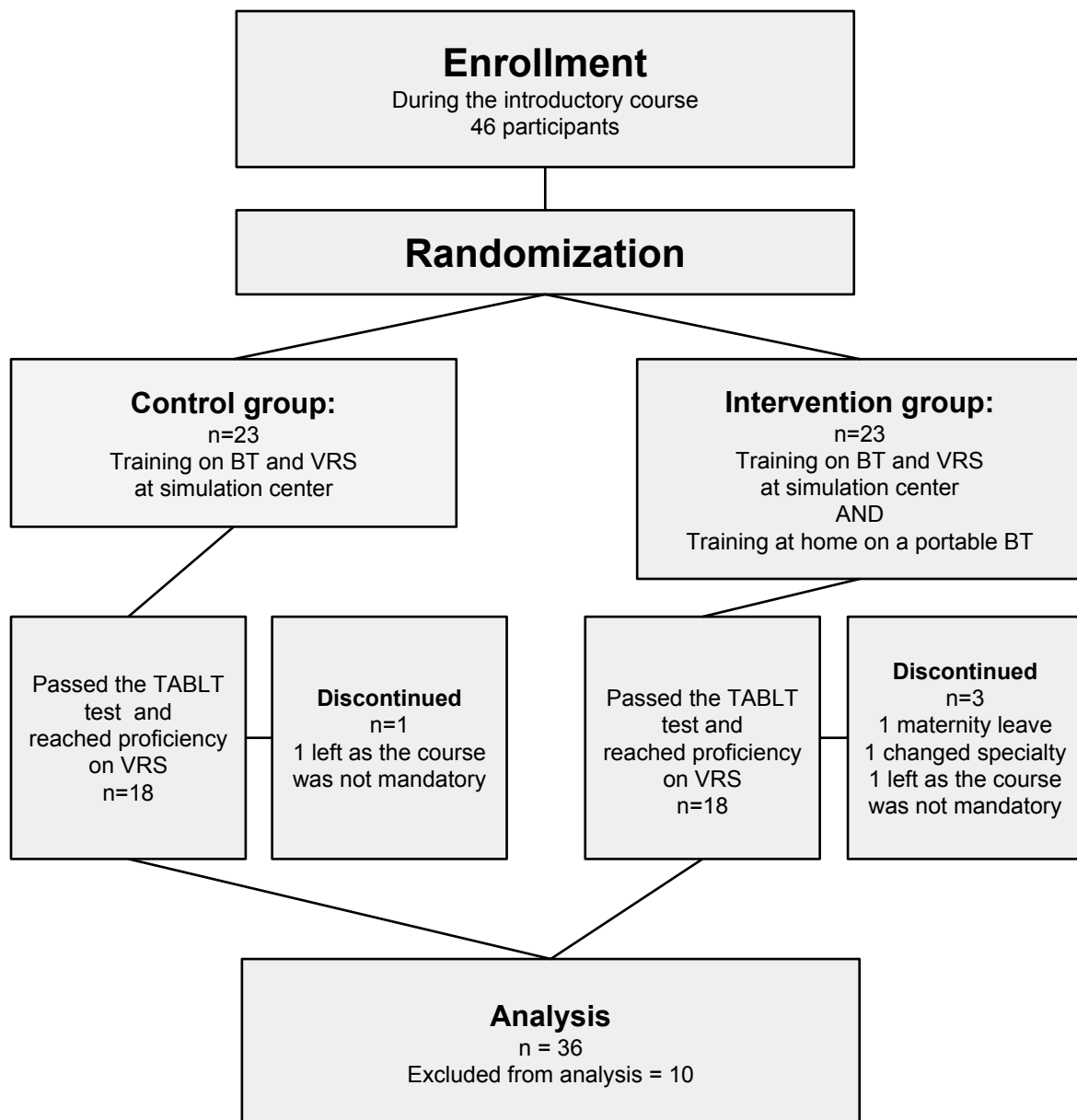
	Intervention group mean (95%CI)	Control group mean (95%CI)	p-value
Number of days	73 (41-106)	76 (41-112)	p = 0.58
Time spent training, in minutes	358 (263-452)	389 (293-485)	p = 0.73
Number of training sessions	3.39 (2.53-4.24)	3.78 (2.9-4.66)	p = 0.89

Figure 1: The Training Program in Basic Laparoscopy

Training Program in Basic Laparoscopy

Introductory course		Self-Directed Training			Operative Course
Theory	Training on BT and VRS		TABL Self-rated Pre-test	TABL Proctored Test	Operating on Sedated Pigs
			Reaching Proficiency on VRS		

Figure 2: Flowchart of Participant Enrollment



TITLE PAGE:

Title: Take-Home Training in a Simulation-Based Laparoscopy Course.

Running head: Take-Home training

Authors:

1: Ebbe Thinggaard, MD, Copenhagen Academy for Medical Education and Simulation, Capital Region Denmark and Department of Surgery, Zealand University Hospital, Koege.

2: Lars Konge, PhD, Copenhagen Academy for Medical Education and Simulation, Capital Region, Denmark.

3: Flemming Bjerrum, MD, Department of Surgery, Zealand University Hospital, Koege.

4: Jeanett Strandbygaard, PhD, Juliane Marie Centre, Rigshospitalet, Copenhagen.

5: Ismail Gögenur, DMSc, Department of Surgery, Zealand University Hospital, Koege.

6: Lene Spanager, Department of Surgery, Nordsjaellands Hospital, Hilleroed,

Author responsible for correspondence and reprints:

Ebbe Thinggaard, Department of Surgery, Zealand University Hospital, Lykkebaekvej 1, 4600 Koege, Denmark, Tel: +45-52722546, ebbe.thinggaard@gmail.com

Conflicts of Interests and Source of Funding:

Equipment was provided by Copenhagen Academy for Medical Education and Simulation, Capital Region, Denmark, Copenhagen, Denmark

Drs Ebbe Thinggaard and Ismail Gögenur report non-restricted educational grants from Johnson & Johnson during the conduct of the submitted work.

Drs Lars Konge, Flemming Bjerrum, Jeanett Strandbygaard and Lene Spanager have no conflicts of interest or financial ties to disclose.

ABSTRACT

Background:

Simulation training can prepare trainees for clinical practice in laparoscopic surgery. Training on boxtrainers allows for simulation training at home, which studies have shown to be a feasible method of training. However, little research has been conducted into how to make it a more efficient method of training. Our aim was to investigate how boxtrainers are used in take-home training to help guide the design of take-home training courses.

Methods:

This study was designed using a mixed methods approach. Junior doctors participating in a laparoscopy curriculum, which included practising at home on box trainers, were invited.

Quantitative data on training patterns was collected from logbooks. Qualitative data on the use of boxtrainers was retrieved from focus groups and individual interviews.

Results:

From logbooks we found that 14 out of 18 junior doctors mixed their training modalities, and four practised first on box trainers then on virtual reality simulators. Twelve practised only at home, while five practised at both places and one practised solely at the simulation centre. After a delayed start, most practised for some time, then had a period without training and then started training again towards the end of the course. We found that the themes of the interviews were: training method, training pattern, feedback and self-regulation. Participants identified the lack of feedback as challenging but described how self-rating provided direction during unsupervised training. Mandatory training elements affected when and how much participants practised.

Conclusions

When participants practised at home, they took an individualised approach to training. They mixed their training at home with training at the simulation centre. Participants practised at the beginning and towards the end of the course. Self-rating helped to guide unsupervised training where feedback was not accessible. Curricular requirements and testing determined when and how much participants practised.

Keywords: Surgery, laparoscopy, boxtrainer, education, training, mixed-method.

INTRODUCTION

Simulation-based training has been widely implemented in laparoscopic training programmes and is now accepted as a supplement to clinical training in laparoscopic skills acquisition. BoxTrainers (BTs) and virtual reality simulators (VRSs) have both been shown to be effective training tools[1]. BTs have the added benefit of being mobile and can therefore help overcome barriers that are associated with training at a simulation centre, such as access, time, and financial constraints[2]. It has been demonstrated that laparoscopic skills can be acquired in unsupervised training at home[3,4]. Training at home allows trainees to decide when they want to train, focus their attention on practising laparoscopic techniques, and avoid practising when they are fatigued[5]. Nonetheless, implementing take-home training can be challenging. BTs are often used minimally by trainees and sometimes not at all in training programmes[6,7]. Some challenges are more pronounced in take-home training, such as the lack of feedback during training. Although studies have shown that training at home is feasible, questions such as how much, when and why trainees use—or do not use—BTs remain unanswered. To examine this, we designed a mixed methods study that aimed to explore the use of mobile box trainers.

MATERIALS AND METHODS

Setting:

A basic laparoscopic training course is offered at the Simulation Centre at Rigshospitalet, Copenhagen Academy for Medical Education and Simulation (CAMES)[8] to doctors in the first year of their specialty training. The course is a cross-specialty training programme for doctors from gynaecology, urology, and surgery departments[9]. The course consists of two formalised course days and a period of self-directed training. The first step is a one-day introductory course that includes theoretical teaching and practical introduction to training on BTs and VRSs. The introductory course is followed by a period of self-directed training, during which trainees can book two-hour training sessions at the simulation centre, where a simulator technician offers technical assistance and feedback. The Training and Assessment of Basic Laparoscopic Techniques (TABLT) test[10] was used for BT training. TABLT is a training and testing system developed for take-home training on BT and consists of five simple tasks: coordination, cutting, sharp dissection, blunt dissection and cyst removal. The TABLT test is supported by evidence of validity and a reasonable pass/fail level has been set. Performance on the TABLT test is measured by using a simple rating system based on time and number of errors[10]. For VRS training, participants practised on the Lapsim® simulator (software version 2014; Surgical Science, Gothenburg, Sweden). The VRS training consisted of seven basic exercises and the bleeding ectopic pregnancy module[11]. As part of the self-directed training, participants have to pass the TABLT test on a BT and reach a predefined level of proficiency on a VRS. From September 2014 to June 2015 we conducted a randomised controlled trial to investigate the added effect of training at home in our existing training programme.

Participants:

Participants were doctors from gynaecology, urology, and surgery departments in their first year of specialty training, who were participating in a randomised trial situated in a basic laparoscopic skills course. Thirty-six participants completed the randomised trial, 18 in each group. Participants from the intervention group (who practised at the simulation centre and were given a mobile boxtrainer[12] to practise with at home) were invited to take part in the study. All participants used a logbook to record their training, including information about when and for how long they practised. Participants were invited by e-mail to take part in focus group interviews and individual interviews at the end of the course. Focus groups interviews and individual interviews were organised to include participants, both men and women, from all three specialties, working at hospitals located at varying distance to the simulation centre. Because the interviews were held at the end of the study, some participants had just finished the course and others had finished the course up to six months earlier. Focus groups were planned so that participants could discuss aspects of box training with one another. Individual interviews were conducted in order to gain an in-depth description and understanding of their training.

Data collection:

First we gathered and analysed the quantitative data, then analysed the qualitative data. Results from quantitative data analysis guided the qualitative analysis. Quantitative data was used for a descriptive analysis of when and how trainees used BT when practising at home. Qualitative data was used to confirm or contradict findings from the quantitative data and to explore the use of BT in more detail. We used focus groups and individual interviews to help improve data triangulation, and to ensure data saturation. Focus groups were used to gain a better understanding of trainee's use of BTs in general, whereas individual interviews were planned to help explore how each individual

structured their training. One interview guideline was developed for the focus group interviews and one interview guideline was developed for the individual interviews. The development was guided by the research aim and the results from the quantitative data analysis. In total, three focus groups interviews and five individual interviews were conducted. The focus group interviews were led by one of the researchers (ET) and one of the other researchers (LS) was also present at one of the focus group interviews. Each focus group interview lasted for approximately an hour and a half and all interviews took place at the simulation centre, a familiar setting for the participants. The focus groups started with an opening exercise in which participants were asked to structure their own ideal laparoscopic training course using pictures of elements from the current training course and clinical practice. The aim of the exercise was to help participants feel comfortable discussing a familiar topic. The exercise was then used as a starting point for further discussion of the use of BT at home. After the focus group interviews, individual interviews were conducted. The guidelines for the individual interviews included 10 questions that had been sent to the participants prior to the interview. Individual interviews lasted approximately half an hour and were conducted at the participant's workplace.

Analysis of data:

For the quantitative data, we conducted a descriptive analysis of when and how the participants used BTs, with the aim of describing training patterns such as duration, number of sessions, and distribution of sessions. Data was presented in the form of graphs that represented the duration of individual and median training durations for the whole group, as well as duration of individual training sessions distributed across the course. The qualitative data analysis was guided by the results from the quantitative analysis. All interviews were transcribed verbatim. Data from focus group interviews and individual interviews were analysed using directed content analysis[13]. In the

interviews each paragraph was paraphrased, and themes were identified from the paraphrased version. When all interviews had been analysed in this way, themes were grouped in main themes and subthemes. Transcripts were then re-analysed using the identified themes to ensure presence in the interviews. Finally, quotes were chosen to help exemplify themes. One focus group interview and two individual interviews were analysed by two members of the research team (LS and ET). Each researcher independently read and analysed the transcripts. Their analyses were then compared and, because there was a strong agreement on which passages to code and on the themes applied, the remaining data were analysed by one researcher (ET). When the analysis was done, quotes were translated in to English.

Ethical considerations:

The trial was submitted for evaluation to the regional ethics committee, which determined that no approval was needed for the trial (H-3-2014-FSP31). The trial was registered at www.clinicaltrials.gov prior to commencing the trial (NCT02243215).

RESULTS

Results from analysis of quantitative data:

Fourteen participants mixed BT and VRS training in no particular order, four participants first practised on the VRS and then on the BT. When training on the BT, none of the participants used the tasks sequentially; all participants changed between tasks. Twelve participants only practised on the BT at home and then came to the simulation centre for the final test; five practised on the BT both at home and at the simulation centre; and one participant only practised at the simulation centre. Training patterns varied substantially and we found the median duration of the training course to be seven weeks (range 6–33 weeks). The median time before starting to practise at home was two weeks (range 1 day to 28 weeks), the median time spent practising was five weeks (range 2–33 weeks), and the median time from passing the TABLT test and to the final operative course was four days (range 1–31 days). For an overview of the individual participants and the median practice duration for all participants, see Figure 1. For an overview of time spent practising, distributed across the training course, see Figure 2.

Results from analysis of qualitative data:

Three focus groups and five individual interviews were conducted. Nine participants took part in the focus group interviews and five participated in the individual interviews; for demographics see Table 1. The following four themes were identified: training methods, training pattern, feedback, and self-regulation. See Table 2 for examples of the analysis.

Training method:

Participants had an individualized approach to training where they mixed the tasks:

“N: I focused on doing the tasks that were already there, so I did not do anything else than those tasks... I did them several times instead of doing all five tasks once and then starting again with all

five tasks.”

N: trainee in surgery

Instead of reaching proficiency on the whole test by doing all tasks in sequence every time, participants practised on each individual task. In general, participants preferred training at home without the interference of peers or supervisors.

“LI: I think I was a little more calm. You could practice in your pyjamas if you felt like it, it was just nice and easy, and if things were problematic it was okay to shout or be annoyed without anyone noticing it.”

LI: trainee in gynaecology

However, one participant preferred training at the simulation centre and did not train at home at all.

“L: I never really got started on training at home... I found that booking a time at the simulation centre was an easier way for me to structure my training. I was sort of forced to be there. When training at home other things always got in the way.”

L: trainee in gynaecology

Training at the simulation centre increased the commitment to training and thereby structured and facilitated training for this person. Nonetheless, most participants agreed that training at home was flexible and easily accessible and therefore preferred training at home:

“P: It is not like you have to go anywhere, and the box trainer, even when you pack it up it does not take up any space. Everyone can have one at home.”

P: trainee in surgery

Training patterns:

Generally, participants felt that starting training at home was a barrier that caused them to delay starting.

“V: It is difficult to say. I think the most difficult part was taking it out the first time... once you got started then things worked out just fine.”

V: trainee in gynaecology

However, a few participants did start just after the introduction course.

“S: ...almost immediately, maybe three to four days after. I thought it was interesting and was like a game you could play. I thought it was fun.”

S: trainee in urology

Nevertheless, most of the participants spent the majority of their training time when the operative course was approaching (and it was a requirement to have passed the test):

“J: The time was a factor in regard to when I started training at home. In the beginning I do not think I actually used it that much. I don’t remember the exact numbers, but when the course was approaching and there was about six weeks until I had to pass the test that is when I started training.”

J: trainee in gynaecology

The participants described a common training pattern in which they tried to be systematic in the beginning but then faced a middle period with low training intensity. They then began to train systematically again when the test was approaching.

“R: Yes, in the beginning it was a bit more systematic, and then there was long period where I did not take it [the BT] out of the bag at all. In the end I practised systematically again, up to this course. At least a couple of times a week.”

R: trainee in surgery

Feedback:

Training at home meant that feedback from supervisors was not available; this was a common

theme that was discussed in the focus groups and individual interviews.

“P: There was one thing missing when you were practised by yourself and that was that you had no supervisors.”

“LI: Yes.”

“S: When you are training in here [at the simulation centre] there was a teacher that could give you some advice. That was missing. But I think the access of training, and being able to train when you feel like it outweighed it.”

P: trainee in surgery

LI: trainee in gynaecology

S: trainee in urology

Feedback was described as being useful but not a prerequisite for training. The increased access to training offered by take-home training outweighed the need for feedback. In fact, some participants considered the lack of feedback to be beneficial since it required them to think for themselves, reflect and take responsibility for their own learning. Participants also discussed how simple exercises may not require feedback.

“N: I also think that if you have a certain level of experience then it may be too much to have someone to instruct you for these tasks. On the other hand, you need to learn how to use your own sense and logic to sort of train by yourself and develop your techniques.”

“N: It requires that you improve by yourself. You do not need supervision for everything and to be told everything, like now you have to take the lead. You have to do it yourself.”

“B: Yes, our consultants always tell us that if you do the same movement three times in a row without progression then you need to change technique.”

“N: Exactly, and if people keep telling you what to do, you end up not knowing how to do anything by yourself.”

N: trainee in surgery

B: trainee in surgery

Another way of receiving feedback was through self-rating, as described by one participant:

“P: On the other one [box trainer], you are giving yourself feedback. You are the one who is rating yourself, there is no one else there.”

P: trainee in surgery

Self-rating was seen as a way to self-evaluate.

“S: Well, you have the option of looking through the video after you have practised. You can actually evaluate yourself.”

S: trainee in urology

Self-regulated learning:

Self-rating was described as a way to regulate one's own learning:

“R: Then you can see what type of errors you made and could keep improving, keep motivating yourself, as now I need to try to avoid this and find out what works and what does not. You try to find your own way, I found that very motivating.”

R: trainee in surgery

Training in this way was motivational and helped increase independence, which not only improved motor learning but also affected cognition. One participant described how self-rating led to more reflection on training.

“N: You reflect more on what was good and what was bad and how you can improve for the next time. It is one thing if somebody else tells you that you have made two mistakes, it is different when you yourself have to remember that it was at this particular time and in that corner of the task that I had a problem.”

N: trainee in surgery

Self-rating and testing during training could provide a clear goal in training.

“P: It helps having specific tasks to train, so you are not just standing there, to have a goal to go for.”

“N: And a score, a rating, something to aim for.”

“P: It actually helps knowing you will be tested. Having to do it properly, and that you have something to work towards.”

N: trainee in surgery

P: trainee in surgery

Testing directed the training and provided the structure and motivation to train. However, training in relation to clinical practice was not largely in focus, although one participant described how training affected his clinical work.

“R: ...when you have practised you feel more confident. Operating the day after [having practised] motivated me to train. Even when you are tired getting home from work, you practice a little more. You could feel the difference at work. When I was practising regularly I felt I really got better [in the OR].”

R: trainee in surgery

In general, participants described how clinical improvement was a motivating factor during the course, although it was not sufficient to maintain training. Testing was the main driver for training.

“N: what makes me train is having to pass a test. Then I will know what to do and it will make me train ...you are lazy when you get home and you do not want to spend your time training.”

“P: Yes, it is the incentive.”

“N: Yes, it is the only thing that motivates you.”

N: trainee in surgery

P: trainee in surgery

DISCUSSION

From the quantitative analysis we found that the participants used an individualized approach to training in which they combined training on BT with training on VRS. The majority practised on a BT at home and on a VRS at the simulation centre. Most participants practiced, in the beginning, after overcoming the initial challenges of setting up the training system. There was then a period with reduced training and then they practised again towards the end of the training programme. From the analyses of the qualitative data we were able to confirm the findings regarding the training pattern and the training method; feedback and self-regulation were identified as additional themes.

A limitation of this study is the sample size for the quantitative part of the study. A larger sample size would have strengthened the external validity of the findings and would also have allowed for a broader sampling for the qualitative study. Sampling was also limited due to the study design, as participants were recruited from a randomized controlled study. Despite these limitations, we still managed to include a wide variety of participants, which helps strengthen the generalizability of our findings. In the qualitative interviews, an outsider could have conducted the interviews in order to help minimise the risk of researchers leading the interviewees to find evidence to support their preconceptions. The qualitative interviews also included participants from different stages of the training programme; some of these participants may have practised a long time ago and could therefore be prone to recall bias. Nonetheless, having participants from different stages of the training programme also meant that some participants had more time to reflect on their training, whereas others would have the training very present. This allowed for a variety of sampling and for different perspectives about how BTs were used in take-home training. Another strength of the study was that we used three sources of data: focus group interviews, individual interviews, and quantitative data from logbooks. This approach helped by providing a rich source of data for data triangulation.

One of the main findings in this study was that the participants used an individualised approach to training. This was evident from the logbooks and confirmed in the qualitative data analysis. That trainees use a more individualized approach when training at home is supported by a randomised study on take-home training[14]. When training at home, participants were free to decide when and how to train, and therefore adopted a more individualised approach to structuring their training and become more self-regulated. We also found that most participants practised on the BT at home and on the VRS at the simulation centre (as VRS was only available there). However, some participants combined training on BT at home and at the simulation centre. Training on a BT both at home and at the simulation centre is a useful method of structuring a simulation-based training programme, as tasks and new training elements can be introduced in on-site training modules[3]. Several studies have compared BTs to VRSs and found both to be effective methods of instruction[1]. However, synergy from training on BTs and VRSs may exist, and mixing the training could be useful. Transfer between BTs and VRSs has been shown previously[15] but the effect of combining the two training methods needs to be explored in more detail. Using different training approaches could be beneficial as it would avoid overtraining and could also help improve basic techniques by spending more overall time on training. When training for specific operations, however, training should be task-specific as this will improve the transfer of skills[16]. Alignment between specific simulated tasks and real operations is essential for preparing trainees for specific operations on patients[17].

At the beginning of the training programme, most participants delayed starting and only started training once they had overcome the initial hurdle of setting up the system. The initial training period was then followed by a period of low training intensity. Towards the end of the training course we found a high-intensity training period as the test approached. This training pattern demonstrates how motivation to train fades when there is no external pressure to do so.

Intrinsic motivation has been shown to be insufficient for trainees to reach a predefined proficiency level[7]. Motivation to train could have been strengthened by facilitating gamification in simulation training. Further studies should look at gamification and its effect on motivation when training at home.

When training at home, the lack of feedback was a major theme discussed in the focus groups and is an inherent challenge in unsupervised training. Feedback has previously been shown to be useful when training complex procedural skills[18]. Nonetheless, participants felt that the easy access to training made up for the lack of feedback, and that the simple nature of the tasks meant that feedback might not have been crucial. One study supports this perception, as feedback improved performance on basic laparoscopic skills but participants were able to complete the training program without feedback[19]. Self-rating was identified as a method of acquiring feedback and helped provide structure in training. One study found that an on-the-fly rating system led to a 100 per cent pass rate on the Fundamentals of Laparoscopic Surgery (FLS) test, which supports the use of self-rating[20]. Self-rating stimulating reflection can lead to self-regulation and can help guide trainees during unsupervised training at home.

Our research supports the use of mandatory training elements as it strongly affects when and how much participants train. A study that introduced mandatory training sessions found that this increased the time participants spent training compared to what has previously been reported in the literature[6]. Another study that supports our findings showed that voluntary practice was insufficient and that obligatory assessments were needed in order to encourage sufficient practice[7]. Participants in our study saw the effect of mandatory training but may have feared the consequences. Perceived consequences could include a later introduction to supervised surgery in the clinical setting, which trainees fear could have reduced the number of operations they would get to do during their training. Therefore, timing of mandatory elements should be planned in relation

to clinical practice in order to help avoid this demotivating factor. Simulation training and clinical practice needs to be combined in a well-structured curriculum to produce optimal results. The best way to combine simulation training and workplace learning in the OR needs to be explored further. Such research could build on theories of communities of practice and situated learning[21] and aim to explore the use of simulation-based training in relation to the progression taking place in workplace learning as the novice moves from observing, to participating as an assistant, to performing supervised surgery, and finally to independent practice.

Conclusions

In this study we found that trainees took an individualized approach to training by mixing training at home and at the simulation centre, both on a BT and a VRS, as they saw fit. Self-rating and testing guided participants in unsupervised training, where it can be difficult to receive feedback. Curricular requirements and testing determine when and how much participants practice.

Disclosures

Drs Thinggaard and Gögenur received non-restricted educational grants from Johnson & Johnson during the conduct of this study.

Drs Bjerrum, Strandbygaard, Konge and Spanager have no conflicts of interest to disclose.

REFERENCES

1. Zendejas B, Brydges R, Hamstra SJ, Cook DA (2013) State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg* 257:586–593.
2. Caban AM, Guido C, Silver M, Rossidis G, Sarosi G, Ben-David K (2013) Use of collapsible box trainer as a module for Resident Education. *J Soc Laparoendo* 17:440–444.
3. Chummun K, Burke JP, O'Sullivan R, Prendiville W (2012) The influence of a 'take home' box trainer on laparoscopic performance for gynaecological surgeons. *Gynecol Surg* 9:303–308.
4. Kobayashi SA, Jamshidi R, O'Sullivan P, Palmer B, Hirose S, Stewart L, Kim EH (2011) Bringing the skills laboratory home: An affordable webcam-based personal trainer for developing laparoscopic skills. *J Surg Educ* 68:105–109.
5. Kahol K, Leyba MJ, Deka M, Deka V, Mayes S, Smith M, Ferrara JJ, Panchanathan S (2008) Effect of fatigue on psychomotor and cognitive skills. *Am J Surg* 195:195–204.
6. Zapf M, Ujiki M (2013) Surgical resident evaluations of portable laparoscopic box trainers. *Surg Innov*. doi:10.1177/1553350614535858, 11 June 2011.
7. van Empel PJ, Verdam MG, Strypet M, van Rijssen LB, Huirne JA, Scheele F, Bonjer HJ, Meijerink WJ (2012) Voluntary autonomous simulator based training in minimally invasive surgery, residents' compliance and reflection. *J Surg Educ* 69:564–570.
8. Konge L, Ringsted C, Bjerrum F, Tolsgaard MG, Bitsch M, Sorensen JL, Schroeder TV (2015) The Simulation Centre at Rigshospitalet, Copenhagen, Denmark. *J Surg Educ* 72:362–365.
9. Bjerrum F, Sorensen JL, Thinggaard E, Strandbygaard J, Konge L (2015) Implementation of a Cross-specialty Training Program in Basic Laparoscopy. *Jsls*. doi:10.4293/JSLS.2015.00059, 2015 Oct-Dec

10. Thinggaard E, Bjerrum F, Strandbygaard J, Gogenur I, Konge L (2015) Validity of a cross-specialty test in basic laparoscopic techniques (TABLT). *Br J Surg* 102:1106–1113.
11. Larsen CR, Grantcharov T, Aggarwal R, Tully A, Sørensen JL, Dalsgaard T, Ottesen B (2006) Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surg Endosc* 20:1460–1466.
12. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P (2013) Tablet-based simulation provides a new solution to accessing laparoscopic skills training. *J Surg Educ* 70:161–163.
13. Hsieh HF, Shannon SE (2005) Three approaches to qualitative content analysis. *Qual Health Res* 15:1277–1288.
14. Korndorffer Jr JR, Bellows CF, Tekian A, Harris IB, Downing SM (2012) Effective home laparoscopic simulation training: A preliminary evaluation of an improved training paradigm. *Am J Surg* 203:1–7.
15. Clevin L, Grantcharov TP (2008) Does box model training improve surgical dexterity and economy of movement during virtual reality laparoscopy? A randomised trial. *Acta Obstet Gynecol Scand* 87:99–103.
16. Bjerrum F, Sørensen JL, Konge L, Rosthoj S, Lindschou J, Ottesen B, Strandbygaard J (2016) Randomized trial to examine procedure-to-procedure transfer in laparoscopic simulator training. *Br J Surg* 103:44–50.
17. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA (2014) Reconsidering fidelity in simulation-based training. *Acad Med* 89:387–392.
18. Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, Gluud C, Grantcharov T, Ottesen B, Sørensen JL (2013) Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized trial. *Ann Surg* 257:839–844.

19. Snyder CW, Vandromme MJ, Tyra SL, Hawn MT (2009) Proficiency-based laparoscopic and endoscopic training with virtual reality simulators: a comparison of proctored and independent approaches. *J Surg Educ* 66:201–207.
20. Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM (2008) Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. *Surg Endosc* 22:1887–1893.
21. Gandamihardja TA (2014) The role of communities of practice in surgical education. *J Surg Educ* 71:645–649.

Tables:

Table 1: Participants Baseline Characteristics:

	Participants using Logbooks	Participants in Focus Group Interviews	Participants in Individual Interviews
No. of participants	18	9	5
Age (median, range)	30(25-36)	31(25-36)	31(28-35)
Gender			
Men	5	5	1
Women	13	4	4
Specialty			
Surgery	6	5	1
Urology	3	2	1
Gynecology	9	2	3

Table 2: Examples of the Analysis

Quotation	Paraphrase	Interpretation	Theme	Sub-theme	Source
<i>“I focused on doing the tasks that were already there, so I did not do anything else than those tasks... I did them several times instead of doing all five tasks once and then starting again with all five tasks.”</i>	Only practiced on included tasks, trained task by task	Participants took an independent approach to training where they mixed the tasks as they saw fit:	Training method:	Individualized training approach	<i>29-yrs old female trainee in surgery</i>
<i>“It is difficult to say. I think the most difficult part was taking it out the first time... once you got started then things work out just fine.”</i>	Unsure, to start training was difficult, and training was easier afterwards	Generally, participants felt that to begin training at home was a barrier and they therefore delayed starting.	Training patterns:	Delayed start of training	<i>28-yrs old female trainee in gynaecology</i>

Figures:

Figure 1: Overview of the training patterns for each participant and median values.

The bars indicate: the time from the beginning of the course until participants started practising at home, the period of active training at home, and the time between passing the test and participating in the final operative course.

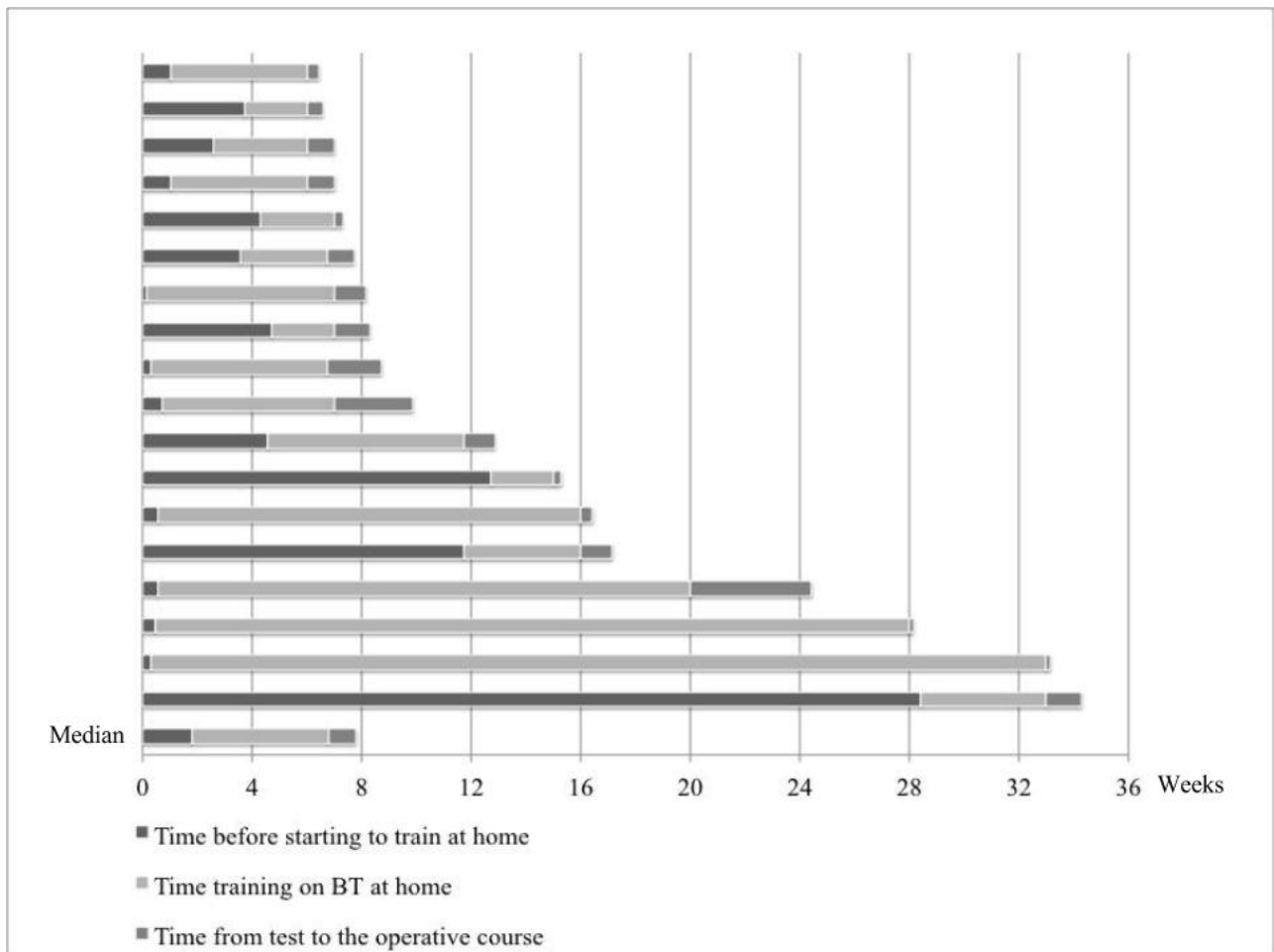


Figure 2: Overview of training: The y-axis shows training time per training session for all participants, and the x-axis shows the relative time passed in the training programme as a percentage.

