Increasing the efficiency of laparoscopic surgical training: assessing the effectiveness of training interventions

Spruit, E.N.

Citation

Spruit, E. N. (2018, January 24). *Increasing the efficiency of laparoscopic surgical training: assessing the effectiveness of training interventions*. Retrieved from https://hdl.handle.net/1887/61139

Author: Spruit, E.N.
Title: Increasing the efficiency of laparoscopic surgical training: assessing the effectiveness of training interventions
Issue Date: 2018-01-24

Increasing the Efficiency of Laparoscopic Surgical Training: Assessing the Effectiveness of Training Interventions

Edward Nicolaas Spruit

ISBN:	978-94-6299-792-9

Omslag:	Sidney Latuperissa
Lay-out:	Nikki Vermeulen | Ridderprint BV
Drukwerk:	Ridderprint BV | www.ridderprint.nl

Increasing the Efficiency of Laparoscopic Surgical Training: Assessing the Effectiveness of Training Interventions

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 24 januari 2018
klokke 13:45 uur

door

Edward Nicolaas Spruit
geboren te Dordrecht
in 1987

Promotor: Prof.dr. J.F. Hamming

Co-promotor: Dr. G.P.H. Band

Promotiecommissie: Prof.dr. F.W. Jansen
Prof.dr. M.P. Schijven (Amsterdam Medisch Centrum)
Prof.dr. T.A.J.M. van Gog (Universiteit Utrecht)
Prof.dr. J. Dankelman (Technische Universiteit Delft)

Contents

CHAPTER 01

INTRODUCTION

Introduction

Over the past two decades, the use of Minimally Invasive Surgery (MIS) has increased substantially in the disciplines of General Surgery, Gynecology and Urology. In practice, MIS has many advantages over traditional open surgery; like reduced damage to bodily tissue, less pain, shorter hospitalization, lower rates of morbidity and mortality, smaller scars and a faster return to normal every-day life (Banta, 1993). For these reasons, patients have demanded this type of surgery more frequently over the past decades, resulting in a wider implementation of the technique. Due to this fast application, research and extensive education on the skills and training required for excellent surgical performance has not yet fully caught up.

The use of a camera, a monitor and laparoscopic tools make MIS operations more complex due to the lack of depth perception, the Fulcrum effect and the reduction in tactile feedback (Hiemstra, Kolkman & Jansen, 2008). As a result of this increase in complexity, different types of skills are required from practitioners in order to ensure safe and efficient surgical performance.

Laparoscopic Surgery

MIS refers to a set of surgical techniques that aim to reduce the size of incisions during surgery. Most MIS procedures are performed in the abdominal cavity (which are called laparoscopic operations), but MIS techniques are also used in the pelvic cavity, in vascular surgery and in other areas like the limbs. During laparoscopic surgical procedures, a few (usually one to five) small incisions are made in the body, thru which a camera and instruments are inserted. The camera (laparoscope) projects to monitors in the operating room visible to the operating surgeon(s) and assisting staff.

Because smaller incisions are made, less damage to bodily tissue is done, typically resulting in less pain for the patient and a faster recovery time. This results in lower chances for infection, morbidity and mortality, shorter hospitalization and better cosmetics (smaller scars). The drawback is that laparoscopic procedures are more difficult to perform than traditional open (laparotomy) procedures. Operating with laparoscopic techniques requires more complex skills and thus makes a procedure more challenging for the operating surgeon.

Challenges of Laparoscopic Surgery

Because laparoscopic procedures make use of small incisions, a camera and monitors, the surgeon and operating team have no direct view of the operating site. As a surgeon is required to look at a monitor to see what is going on, his hand-eye coordination is altered substantially. Also, the captured image of the laparoscope is enlarged several times resulting in a transformation of visual-motor proportions. This means that a very fine motor movement of the hand on an operating instrument results in a much larger movement of the corresponding instrument on the monitor screen.

Another limitation of viewing the operating site from a monitor screen is that the surgeon's visual perception is limited to a two-dimensional video image. This hampers the estimation of depth relations in the operating field. In traditional open surgery the surgeon can use both eyes to directly look into the body of a patient, but in laparoscopic surgery (LS) this default perception of depth (called stereopsis) is lost. Furthermore, the laparoscope can only provide one perspective (at a time) of the inside of the abdominal cavity, depending on the angle and position that it is being placed. This implies that the surgeon has to cope with an incomplete image of the operating site, which increases chances of unnoticed injury occurring outside of the visual field.

In addition to these perceptual constraints, LS makes use of long instruments and surgeons cannot use their hands directly to operate. In traditional open surgery, the operating surgeon can feel the texture and warmth of different bodily structures and make judgments and decisions about anatomy accordingly. In LS however, this degree of haptic sensation is lost, as laparoscopic instruments provide very little tactile feedback. Also, the instruments tilt around an axis (the abdominal wall), resulting in inversed movements. Moving the handle of an instrument to the left, results in a movement to the right of the tip of the instrument inside the body of the patient (and vice-versa for left to right and up and down). This is called the Fulcrum Effect (Gallagher, McClure, McGuigan, Ritchie & Sheehy, 2008); since it is the same inverse movement that pilots have to face in handling a control stick in an airplane (the base of a control stick is called the fulcrum). This makes the handling of laparoscopic instruments very counter-intuitive and it requires more practice to become proficient at it. All of these challenges make anatomical orienting and mobilization of organs and bodily tissues during an operation a lot more difficult.

Teaching Laparoscopic Surgery

Aside from bearing a larger burden on the cognitive skills of the surgeon and operating team, these challenges also open a window of risk for the patient. In the hands of an under-skilled surgeon, the choice to opt for a laparoscopic procedure may actually result in more damage to bodily tissue, rather than less. Jordan, Gallagher, McGuigan and McClure (2000) suggest that experience of surgeons and complication rates are associated. The authors highlight the findings of Wherry et. al. (1994), pointing out that most complications occur in the first ten laparoscopic procedures a surgeon performs. They state that: “it is important to identify scientifically the problems that face the trainee laparoscopic surgeon, and also to investigate strategies to overcome these difficulties.” (Jordan et. al., 2000).

Reduction of damage to bodily tissue is one of the most important drives behind the development of MIS (Hamming, 2005), which makes the importance of proficient surgeons apparent.

Surgical procedures are traditionally taught to residents in a Halstedian mentor-apprenticeship model (Gallagher & O’Sullivan, 2012). In this scenario, residents are allowed to perform steps in an operation under the direct supervision of an experienced surgeon. In the first years of training, the resident mainly assists in basic surgical procedures and performs relatively easier tasks during an operation. As experience and skill accumulate, residents learn to perform more difficult steps in surgery.

Since the introduction of LS in the late eighties, the necessity of alternative training methods has become apparent, as the skills required for LS are more difficult to teach in the OR. Also, public awareness regarding the safety of surgery had altered substantially as a result of iatrogenic errors reported in popular media and authority assessments (Gallagher & O’Sullivan, 2012; van der Wal, 2007). The combination of these factors led to the questioning of the traditional mentor-apprentice model of teaching surgical skills. Most surgeons will agree that there is no substitute for operating on real patients in learning to perform surgery. There are, however, several drawbacks to this model and trainees can be taught specific laparoscopic skills through simulation training.

First of all, it is not ethically appropriate to let an inexperienced surgeon operate on patients, even in the presence of a mentor. The equality of health care distribution cannot be guaranteed when some patients receive treatment in a teaching hospital, whereas some patients are operated on by experienced surgeons. To ensure quality of health care, sufficient skills need to be mastered by residents before the first contact with patients.

Second, residents are working fewer hours in a week than before, which minimizes the amount of time that is available for teaching (Reznick & MacRae, 2006). This lack of time allocated to training may pose problem to learning efficiency. Therefore, training surgical skills in the traditional model may not be ideal. In the OR, providing quality health care to the is the primary goal (i.e. a successful surgical procedure with no complications and a positive outcome for the patient). The residents’ learning process will always be secondary to providing quality health care, making it a less than optimal learning environment. In these circumstances, the teaching process of a surgical resident will always be subjugated to the safe, fast, and successful completion of the operation. In contrast, in a training that is primarily geared towards teaching a resident key knowledge and skills, the learning process of is the primary goal and the trainee determines the speed and complexity level to train at.

Third, the mentor-apprenticeship model creates a very diverse training program for different surgical residents. What is learned is dependent on the case-load and preferences of a specific proctoring surgeon, so that all residents receive different curriculum content during their training. Proficiency-based training on the other hand, focuses on identifying the important skills needed for excellent performance in LS and measuring and documenting the level of those skills (Aggarwal & Darzi, 2006). A more evidence-based approach to the design of curriculum is desirable (Levinson 2010), for surgical training, as well as medical education in general.

In the Netherlands, the risks of MIS have received more attention from the Dutch government, highlighting the incident rates of complications during LS (van der Wal, 2007). In 2008, the inspection of Health Care asked all Dutch hospitals to write a plan of action for improving safety and competence of LS procedures.

In response to this message, medical associations and institutions (NVEC, 2009) have created more guidelines for MIS and their attention has turned towards finding ways to bring more structure and efficiency to the training of laparoscopic skills to medical residents. On an international level, the need for new kinds of training for surgery was stimulated by a series of reports of malpractice spread across the globe (Gallagher & O'Sullivan, 2012). Media coverage of some of these scandals led to a demand for more rigorous assessment and qualifications of surgical skills. As a result of this, the traditional mentor-apprentice model of learning for the resident surgeon was put into question. Surgeons found that their residents had more difficulty learning laparoscopic skills as the acquisition of these surgical skills is a more complex endeavor. These challenges of learning new laparoscopic procedures, ethical issues, along with a reduction in work hours facilitated the need for a new kind of training. Ultimately, all of these factors combined led to the introduction of simulation training in surgery.

Optimizing Efficiency of Laparoscopy Simulation Training

Simulation training of surgical skills must be researched thoroughly, in order to be valid (predictive of later surgical performance). By taking into account the perspectives of multiple disciplines (surgery, educational, cognitive and organizational psychology); ways can be found in order to enhance laparoscopic skills training.

A way in which laparoscopic skills training may be improved is by exploring what factors influence surgical performance, how laparoscopic skills are learned most effectively, how these surgical skills are best maintained over time and how the acquisition of skills differs among individuals as well as training conditions.

The aim of the research project described in this thesis was to study and improve the training of laparoscopic skills for surgeons, in order to increase patient safety in the operating room.

The goal of this project was to take a good look at these research questions by investigating the elements that influence the practice of laparoscopy. The project will be a combination of qualitative and quantitative methods, including a cognitive task analysis and several lab studies that will focus on the following research question:

- Under which training conditions (schedules, training design, feedback, etc.) are laparoscopic skills most efficiently learned and retained on the long-term?

The results and outcomes of these studies will serve as indicators of how current training and curriculum may be modified and optimized. The goal is to facilitate trainers and instructors in medical centers in their decisions regarding training design.

In the studies described, we draw on concepts from educational and cognitive psychology, that have been viable principles to improve learning efficiency and long-term retention in other contexts (knowledge acquisition, basic motor tasks, aviation, sports, medical education). In the chapters to come, we also describe the cognitive capacities inherent to the learner, which are an important part of understand why a learning principle works and may prove useful when considering the use of learning principles in training.

Chapter Overview

In order to provide an accurate picture of what the practice of LS entails, a cognitive task analysis was conducted. This analysis included reviewing literature, analyzing recorded video footage of laparoscopic procedures, observing operations in the OR and conducting semi-structured interviews with experienced laparoscopic surgeons. The results of this analysis can be found in **Chapter 2**. This chapter gives the reader an impression of what an advanced laparoscopic procedure entails and gives a brief summary of all the important aspects for training surgical residents.

The dissertation then continues, zooming in on one of the aspects of training, namely the perceptual and motor skills required to perform laparoscopic surgery. The next **Chapter (3)** is an in-depth literature review into the possible interventions which can be applied to laparoscopic motor skill training, which serves as a basis for the empirical lab studies that follow in later chapters.

Chapter 4 is a brief commentary on the topic of how instructor feedback influences a trainee's learning process. **Chapter 5** continues on the topic of feedback, but on the level of the training simulator. It describes a study with the aim of finding the optimal dosage of presenting visual force feedback to trainees during training.

In **Chapter 6**, a varied practice intervention was tested, by interleaving different variants of laparoscopic training tasks. Also we used alternating angles of the laparoscope during training, since the angle of the laparoscopic is not always constant during LS procedures in the OR. **Chapter 7** describes a study focused on comparing two groups with different training time schedules (spacing) in order to assess the viability of (longer) time intervals between training sessions.

In **Chapter 8**, we revisit the topic of spacing training, so multiple groups with different training schedules are compared. In this study, we aim to specify the influence of breaks, fatigue and sleep consolidation on the effectiveness of spacing training. The final chapter consists of a general discussion with recommendations. A final disclaimer is that this dissertation was written as a collaborative research project between medical and psychological departments. This reflects in the writing with some chapters (and abstracts) being more concise and to the point (medical style, chapter 4, 5, 7 and 8) and some being more theoretical, descriptive, elaborate and with an overall more thorough exploration of the existing literature (psychological style, chapter 2, 3 and 6).

CHAPTER 02

COGNITIVE TASK ANALYSIS OF LAPAROSCOPIC SIGMOID COLON RESECTION: TOWARDS A STANDARDIZED CURRICULUM FOR TRAINING LAPAROSCOPIC SURGERY

Edward N. Spruit MSc.
Tom Hulscher MSc.
Dr. Guido P.H. Band
Prof. Jaap F. Hamming

Abstract

Background: The complexity of laparoscopy has raised the urgency for standardizing the associated curriculum. We used cognitive task analysis (CTA) on laparoscopic sigmoid colon resection to identify essential action steps, decision points and key skills.

Methods: A cognitive task analysis was performed using literature, interviews, video, and observations in the OR of laparoscopic procedures. Video-recorded semi-structured interviews were held with seven experienced surgeons from various medical centers in the Netherlands.

Results: A procedural checklist was established that contains seven sub-tasks, 41 operative steps and seven decision points. The difficulty of each operative step and decision point was rated (novice, intermediate, advanced) by a panel of experienced surgeons. The CTA was used to create a surgical skill profile that can guide the design of a standardized curriculum for training residents.

Conclusions: The surgical skill profile and procedural checklist can serve as a framework for standardized training and assessment of laparoscopic skills.

Introduction

Laparoscopic surgery (LS) differs substantially from open forms of surgery in its cognitive demands and the complexity of perceptual and motor skills (Gallagher & O'Sullivan, 2011). For example, instruments can only reach the organs through trocars in the abdomen, which limit the freedom to move, while vision is limited by the orientation of a camera and the two-dimensional display. Therefore, preparing surgical residents for their first LS on patients requires efficient and thorough training based on a detailed understanding of the challenges involved. It is in the interest of both patient safety and the expenses of medical training that all residents follow a well-defined, optimized and evidence-based curriculum based on principles derived from cognitive and education science (Friedlander et al., 2011; Hodges & Kuper, 2012; Levinson, 2010).

As an important step towards standardizing the curriculum of residents acquiring LS skills, a cognitive task analysis (CTA) of LS is reported. This method has been used effectively in the past to analyze colonoscopy (Sullivan et al., 2008), laparoscopic appendectomy (Smink et al., 2012) and laparoscopic Nissen fundoplication (Peyre et al., 2009). In this paper we applied CTA to analyze advanced colon surgery. The CTA results consist of a step and decision checklist for laparoscopic sigmoid colon resection and a surgical skill profile that can serve as a guide for designing the associated training curriculum.

Method

Cognitive Task Analysis (CTA)

In order to establish a standardized curriculum for LS, we performed a CTA to chart which specific skills residents need for LS, and what they should be trained in. It included reviewing literature, analyzing recorded videos of laparoscopic procedures, observing operations in the OR and conducting video-recorded semi-structured interviews with seven experienced laparoscopic surgeons from several medical centers in the Netherlands. Automatized performance following years of experience often prevents experts from articulating all the required concepts, cues, action steps and knowledge, because the task has become second-nature to them (Clark, Feldon, van Merriënboer, Yates & Early, 2008). CTA provides research techniques that can be used to elicit this covert knowledge and capture the cognitive decision-making processes that experts utilize while performing a task. This can be done by asking probing questions and process tracing in a semi-structured interview (Clark et al., 2008; Militello & Hutton, 1998). In this way, CTA can help to capture the goals, steps, decision points, skills, knowledge and demands that are essential to the safe and successful completion of a laparoscopic operation.

The CTA was applied to three types of LS: laparoscopic cholecystectomy, laparoscopic fundoplication and laparoscopic sigmoid colon resection. For this paper, we focused on sigmoid colon resection in depth and will use it as the exemplary case of LS. Although most tasks during LS are performed sequentially, as summarized in the procedural checklist in Appendix A, additional tasks continue throughout the operation, such as team communication, monitoring the operating field, and adjusting the pace of the procedure. Based on the CTA, we established a set of key skills for LS, which can be found in Appendix B. The first part of the article covers the analysis of laparoscopic sigmoid colon resection. The second part describes a skills profile of what makes for a proficient laparoscopic surgeon and offers brief suggestions for how these skills may be integrated into a curriculum for surgical residents.

Laparoscopic Sigmoid Colon Resection

Laparoscopic colon surgery is an advanced form of surgery, which a surgical resident learns to perform in later stages of training after basic skills and procedures have been mastered. In this analysis, we focused specifically on laparoscopic sigmoid colon resection for the treatment of cancer.

Operative Steps and Decision Points

From the Cognitive Task Analysis, we were able to divide the operation into seven sub-tasks with 39 operative steps of which six are optional or dependent on the given conditions. Seven decision points were identified for the procedure. The full list of operative steps and decision points can be found in Appendix A.

The procedure starts with patient positioning and trocar placement, a sub-task that is very similar for most laparoscopic procedures. After this has been established, the surgeon investigates the abdominal cavity for potential abnormality of other organs or the peritoneum and accesses the local tumor status. In the case of advanced local tumor growth, the surgeon may opt to convert to laparotomy. In the majority of cases, a medial to lateral approach is used for the sigmoid dissection.

Blunt dissection in the avascular layers of the sigmoid mesocolon is initiated to create a submesentaric window (tunnel) in order to expose sensitive structures like the left ureter and the hypo gastric nerve. When these have been identified, the artery supply that supplies the sigmoid colon is dissected first. This ‘vessels-first’ approach is a common oncological principle to limit potential tumor cell dissemination. The sigmoid arteries need to be dissected in proximal direction so that at least twelve lymph nodes are attached to the extracted colon specimen.

In the sub-task that follows, the posterior and lateral attachments of the sigmoid are mobilized. If there is a tension-free fit of the proximal and distal end of the remaining colon for anastomosis, no further mobilization is needed. If there is not enough slack however, the descending colon and splenic flexure must be mobilized to create more length on the colon. The need for further mobilization of the colon varies among patients and is also influenced by the location of the tumor. If more length is needed, the descending colon is mobilized laterally at first and medially if necessary. Some surgeons always mobilize the descending colon as a protocol.

When a tension-free fit is achieved, the colon can be divided distally at a minimal margin of five centimeters below the tumor. After this, the proximal colon is extracted outside the abdomen, divided proximally five centimeters above the tumor while also taking into account the artery supply.

In the final sub-task, the anastomosis of the proximal and distal colon is created. If desirable, a deviating ilieostoma is created and the anastomosis is tested for leakage. If necessary, the mesocolon is closed to prevent a potential herniation.

Anatomical landmarks and hazard zones

The most important structures to preserve in this operation are the left ureter and the hypogastric nerves. In case of upward mobilization, the spleen and pancreas are anatomical landmarks. Also, care must be taken not to damage the colon itself. Furthermore, correct identification of the inferior mesenteric artery and its three most important branches (the arteria colica sinistra, the trunk of the sigmoid arteries and the arteria rectus superior) are very important.

There is the choice between doing a high-tie or low-tie ligation of the blood supply to the colon. A high-tie implies dissecting the inferior mesenteric artery and thereby take out all its branches, whereas a low-tie implies only ligating those arteries that supply the part of the colon that is going to be extracted (Lange, Buunen, van de Velde & Lange, 2008). For optimal vascularization of the remaining colon, the low-tie procedure is preferred.

Individual case variability

There are anatomical variations between patients. For example, patients from developed countries who eat a lot of red meat tend to have a lengthier colon, which makes the necessity of upward mobilization less likely. Also, the procedure tends to be facilitated in lean patients who have a low-level of intra-abdominal fat. In these patients the arteries of the sigmoid mesocolon are much easier to identify. For the same reason, the operation tends to be easier to perform on obese women, as female patients usually have more subcutaneous rather than intra-abdominal fat in comparison to obese men. When the medial mobilization and ligation of the arteries becomes too troublesome, a surgeon may opt to use a lateral-to-medial approach in difficult cases. Awareness of individual differences is important to teach to surgical residents.

Surgical Skills Profile

One of the main incentives for the introduction of LS has been to limit damage to bodily tissue (Hamming , 2005), while still achieving all objectives of surgery. The combination of these two goals defines surgical proficiency. A more detailed profile of surgical proficiency was derived from the CTA. An overview of this skill profile is displayed in Appendix B. Each point of this proficiency profile and its place in training will be discussed briefly in the remainder of this article.

Surgical Knowledge

Surgical knowledge consists of semantic knowledge regarding anatomy, pathology and procedural steps. Semantic knowledge refers to general knowledge that has validity across different situations. It is typically acquired through books, lectures and educational videos. Knowledge of anatomy and physiology is an obvious prerequisite for LS. Insufficient knowledge could lead to incorrect decisions and errors with dramatic results, as the surgeon needs to know where incisions can and cannot be made safely. Valuable suggestions for the anatomy curriculum have been made in literature (Kooloos, de Waal Malefijt, Ruiter & Vorstenbosch; Sugand, Abrahams & Khurana, 2010).

Regarding knowledge acquisition, it is notable that educational psychology suggests ways to optimize learning, which are not yet commonly implemented in medical curricula. For example, prolonged retention of knowledge is achieved by distributing the same study time over an extended period and by interleaving different types of practice problems (Dunlosky, Rawson, Marsh, Nathan & Willingham, 2013; Rohrer & Pashler, 2010). However, the medical curriculum typically works in dedicated thematic blocks (minimal spacing and interleaving of study material), which is suboptimal for learning efficiency.

Declarative surgical knowledge, that is, explicit knowledge of the steps to go through for each LS type, is relatively easy to acquire, even without practical training. Furthermore, because of the consistency inherent in most LS protocols, retrieving the order of steps from declarative memory can be easily trained to perfection. In contrast, procedural knowledge, that is memory for actions that cannot be articulated (such as LS motor, perceptual and clinical decision making skills), has a shallow learning curve.

Acquiring procedural knowledge is facilitated by using worked examples: initially, the structure of a problem to solve can be simplified by step-by-step demonstrations (van Merriënboer & Sweller, 2010). Because worked examples involve a modest cognitive load during training, residual processing capacity remains available for the actual learning process, rather than all being used for understanding one's own actions or instructions. Likewise, for complex tasks, there is a benefit to first learning parts of a task before learning the task as a whole, provided that the elements learned are independent (Teague, Gittelman & Park, 1994; Wightman & Lintern, 1985). This experimental finding argues against immersion of students into the full complexity of LS in the starting phase of their training (Spruit, Band, Hamming & Ridderinkhof, 2014, see chapter 3).

Perceptual skills

During surgery, various types of tissue respond differently to being dissected. Some tissues are avascular and can be dissected easily, while others will bleed more readily. Many of our experts stated that an operation is much easier to perform if one stays in the proper planes and areas during dissection. A clear example of this is when surgeons are creating the submesentaric window during laparoscopic sigmoid colon resection or dissection in the Total Mesorectal Excision (TME) plane during rectal resections. Because of this, being able to identify the different types of tissue is crucial in completing a surgical procedure and minimizing damage. This perceptual skill is the result of many years of experience in the operating room, although it can also be trained in anatomy labs.

Technology is currently in development that can distinguish between different types of tissue (mesenteric fat, blood vessels, colon, ureter, tumor) and can be implemented as an image enhancement aid during colorectal surgery (Schols, Dunias, Wieringa & Stassen, 2013). Similarly, our panel of surgeons stated that color marking techniques are already used in order to spot the tumor more easily during colon resection. There is a realistic expectation that technology will facilitate more accurate identification of tissue in the future.

Additionally, there is potential for e-learning for acquiring the skill of identifying different tissues. For example, recorded footage of laparoscopic procedures may serve as a medium for teaching surgical residents how to identify different kinds of tissue and anatomy, if the footage is edited and highlighted in a way that facilitates the learning of this skill.

The ability to form a clear mental picture of the kind of action to be performed is indispensable to a surgeon. This skill (Jeannerod, 1995) involves creating a mental image of the end state that is to be achieved as well as a mental representation of the motor action required to achieve that end state. A practical example of this skill is knowing exactly what the optimal angle is to grasp a needle so one can set it up perfectly for making throws during laparoscopic suturing. With practice, the surgeon will acquire enough reference experiences to immediately retrieve the most efficient action representation to grasp the needle optimally for a smooth completion of the task (Logan, 1988).

Adapting to reduced tactile feedback and depth cues occurs during basic laparoscopic skill courses using Virtual Reality (VR) - and physical box-trainer simulators or in porcine models. A common complaint about VR-simulators is that tactile feedback is lacking and even for the ones that do provide it, surgeons often perceive the feedback to be unrealistic (Våpenstad, Hofstad, Langø, Mårvik & Chmarra, 2013). Recently, measures have been taken to improve tactile feedback on instruments in a physical box trainer, which can lead to positive effects on grip force and tissue handling (Horeman, Rodrigues, van den Dobbelsteen, Jansen & Dankelman, 2012; Wottawa et al., 2013).

Motor skills

A cornerstone of surgical proficiency in laparoscopy is motor skill, since the used instruments are different from open surgery. Adapting to the fulcrum effect (the fact that instruments move inversely in- and outside the abdominal cavity) and familiarity with the instruments form the basis of motor skill.

The correct application of motor skills is highly intertwined with the surgical knowledge a resident has and the quality of the (processed) perceptual input that is coming in. Hence, proficiency in motor skills, though essential, is no guarantee of adequate application of those skills during surgery. Without sufficient practice, such skills decrease rapidly (Gallagher, Jordan-Black & O'Sullivan, 2012). This has been reason for the Dutch endoscopy society to prohibit infrequent LS performance (Bemelman, 2009).

Because the images coming from the laparoscope are magnified on a screen, the motions a surgeon makes with instruments are also magnified, leading to amplified tremor, another challenge a resident needs to master. Surgical accuracy can be improved with motion scaling by means of robotic surgical systems (Prasad et al., 2004), although its application in most hospitals is limited due to high expenses of surgical robots and also because it is still unclear whether or not the use of robots in the OR will lead to superior outcomes for patients in comparison to conventional techniques (Wiklund, 2004), a common issue with the introduction of new technology in healthcare.

Basic motor actions include mobilizing tissue for exposure and dissection, blunt and sharp dissection, use of devices (cautery, ultrasonic shears, LigaSure, stapler), intra-corporeal suturing and additional techniques (such as the use of umbilical tape, meshes and extraction bags in various procedures). Most of these motor skills can be adequately trained during specific courses on LS and in most teaching hospitals that have a skills lab with laparoscopy training setups like physical box-trainers, virtual reality simulators or porcine models (Aggarwal, Moorthy & Darzi, 2004).

For optimal training design of a laparoscopic motor skills program, factors such as proficiency benchmarks, spacing, adaptive training, task variability, part-task training, mental imagery, dual-task training and goal-directed deliberate practice should be taken into account (see Spruit et al., 2014 (chapter 3) for an in-depth review).

Clinical decision making

In medicine, physicians have to make important decisions about choosing or not choosing different treatments on the basis of a patient's symptoms, the outcome of diagnostic tests and probabilities (Kassirer, 1976). During surgical procedures, this decision making is expressed in the art of knowing how and when to apply certain skills in order to facilitate a positive outcome for the patient. It is knowing when to switch to another task or when to inhibit an automatic response when an unusual case presents itself. This form of proficiency can be viewed as a combination of faculties that determine what stage the operation is at, monitor the perceptual input that is coming in, and pick a course of action to proceed with accordingly. In this way, the procedure is directed to meet the two primary goals (achieving the objectives of the operation, while minimizing damage) in the most successful manner that is possible given the current level of the surgeon's proficiency and the current external conditions. According to our panel of experts, proficiency in this skill is the result of years of practical experience. When humans are engaged in a task, they tend to make decisions intuitively in-the-moment and reflect on them after completing the task. Especially in emergency situations, surgeons need to be able make snap judgments on what to do in order to respond fast and accurate. There are exceptions to this, as was reported by surgeons in our interviews. In some circumstances, a surgeon may opt to take a short break in the operation to consciously reflect on the next course of action. For example, when a procedure progresses differently from what was initially expected, it may be a wise decision to stop and pause in order to think thoroughly about how to adjust the plans for the rest of the operation. Nevertheless, the actual decision to know when to stop and take a short break still classifies as a form of intuitive clinical decision making that comes with experience. The ability to know when and how to regulate conscious attention is an example of clinical decision making. The actual regulation of attention itself is a process that influences the outcome of a procedure, as well as the learning rates of surgical residents. If residents can recognize when their attention falters and a break would be beneficial, it will increase their learning efficiency. A good surgical curriculum ought to incorporate meta-cognition (learn residents how to learn efficiently).

Automatizing motor behavior is both a key and a threat to safe surgery. While in general, automatization of skills is crucial to reduce the sensitivity to stressors, it also hampers flexible task performance under unusual circumstances. As a consequence, the necessarily deviant execution of a standard operation introduces the risk of falling back into habitual execution. Knowing when and how to adjust is the result of years of experience, but the foundation for both fluency and flexibility should be fostered in surgical residents. Residents appreciate the inclusion of complex cases for decision making into the curriculum (Cook, Beckman, Thomas & Thompson, 2008). Learning about unique cases during training courses may lead to better understanding of individual variability in patients and enhanced clinical decision making. In addition, flexibility can be achieved by interleaving training tasks (McDaniel, 2012) and introducing variable motor conditions of practice such as mirroring the image display (Jordan, Gallagher, McGuigan, McGlade & McClure, 2000).

During acquisition of new skills, the burden on the resident's mental resources such as working memory is still high. This has implications for the way LS can be taught. That is, learning is hampered by teaching too many new skills at once or by presenting more information than the bare essentials for learning (van Merriënboer & Sweller, 2010). Since performing an entire laparoscopic operation would be too overwhelming for a novice, a trainee needs to demonstrate proficiency on different simpler training tasks first. As these basic training skills are practiced, these will become automatized over time, meaning that executing them will require less attention. As efficiency of basic surgical skills increases, more spare attention becomes available, which in turn allow the surgeon to handle the cognitive demands of a full operation (Gallagher & O'Sullivan, 2011). Thus, trained surgeons are increasingly able to detect anomalies, inhibit habitual actions, explore alternatives, and switch to an alternative approach, despite the intrinsic cognitive load of LS. This makes it less likely for a more trained surgeon to become overloaded when something goes wrong during an operation, as there will be enough attention available to deal with a possible emergency.

Additionally, a strong focus on career-long learning and deliberate practice (Ericsson, 2006) should be included in the curriculum. Our panel of surgeons reported that a sense of perfectionism is very important in their line of work and deliberate practice is characterized by a strong impulse to continue to master skills long after the point where proficiency has been reached during the period of training.

Mental endurance

Another important factor while performing surgery is the ability to focus on a task for extended periods of time, during monotone and routine steps as well as in cases of emergency. As it is important for a surgeon to develop the ability to stay calm and focused during emergencies, emotional stability is also an important psychological factor for predicting laparoscopic performance. The skill to stay emotionally centered during times of adversity in an operation is not limited to the degree of automation of basic surgical skills; it may also be influenced by other factors. Our panel of surgeons reported that emotional stability can be influenced by social, health and personality factors. If a surgeon is not well-rested, physically fit, or has an unresolved issue in his or her social life, this may have a detrimental effect on their performance. All of these may result in a compromised emotional stability in the operating room.

Also, residents should be taught about healthy posture, ergonomics (Van Det, Meijerink, Hoff, Totte, & Pierie, 2009; Wong, Smith & Crowe, 2010) and drawing healthy personal boundaries, in order to minimize fatigue, injuries or burn-out. Research has shown that performance on a surgical simulator deteriorates after working night shifts (Leff et al., 2008) and research on cognitive tasks has repeatedly shown that performance and especially cognitive control (Lorist, Boksem & Ridderinkhof, 2005) is impaired after inducement of mental fatigue. In order to ensure patient safety and to lower chances of burn-out, the surgical curriculum should emphasize the causes, symptoms and consequences of mental fatigue and encourage surgeons to develop healthy working schedules and conditions (Balch, Freischlag & Shanafelt, 2009).

Social skills

During advanced laparoscopic procedures four or more trocars are used, which means that an assistant is always required for handling the camera and usually a grasper instrument. Similarly, there are more OR team members who assist in preparing all the required instruments and materials. Hence, the ability to co-operate effectively with team members is essential, as complications may arise during surgery that requires quick teamwork to respond to effectively. The primary surgeon should be able to instruct all the assisting staff successfully, even when these are inexperienced. Also, surgeons should be aware of hierarchies in the team (Undre, Sevdalis, Healey, Darzi & Vincent, 2007) and avoid blocking out valuable feedback from assisting team members.

Social skills can be trained by designing training formats that involve residents and co-assistants to practice skills in co-operation. An example of this is medical team training, in which the focus is placed on communication skills between surgeons, nurses and anesthesiologists based on principles of crew resource management (Awad et al., 2005). Also, learning rates can be improved by systematically implementing mutual feedback practices as a standard protocol after each procedure (London & Beatty, 1993). OR team members should also be educated regarding the nature of communication failures in the OR (Linhard et al., 2004) so that they can adjust their future behavior to avoid them.

Technology skills

According to our panel of surgeons, affinity with the hardware of laparoscopic surgery is another important skill. As one can imagine, being able to trouble-shoot a coagulation device or an ambiguous monitor image is highly preferred over calling in a technician during a surgical procedure. In this area there is potential for much improvement, as not all current forms of training familiarize students with technical possibilities (like contrast or shutter speed settings for example).

It is not uncommon for medical centers to purchase new technology without providing adequate training to the staff members using it. Training courses that teach the ins and outs of OR hardware would be of great benefit to the curriculum for surgical residents. E-learning courses on technology, using multimedia (Issa et al., 2011) and cognitive load (van Gog & Paas, 2008) principles can be used to train these skills.

Conclusions

In order to explore what constitutes expertise in LS, a CTA was conducted. The CTA yielded two results: (1) a procedural checklist for Laparoscopic Sigmoid Colon Resection with recommendations of difficulty levels for each step; and (2) a surgical skills profile that can be used as a guideline for creating a standardized curriculum for teaching residents in LS. It is desirable that each surgeon's proficiency on important skills for LS can be charted and documented accurately in order to ensure valid self-efficacy of skill of surgeons, quality health care and transparency to the general public.

Acknowledgements

Dr. Bonsing, LUMC; Dr. Hueting, Diaconessenhuis Leiden; Dr. Mannaerts; Sint Franciscus Gasthuis Rotterdam; Dr. Tanis, Academic Medical Centre Amsterdam; Dr. Kimenai, Erasmus Medical Centre; Dr. Lamme and Dr. Hesp, Albert Schweitzer Ziekenhuis.

CHAPTER 03

OPTIMAL TRAINING DESIGN FOR PROCEDURAL MOTOR SKILLS: A REVIEW AND APPLICATION TO LAPAROSCOPIC SURGERY

Edward N. Spruit MSc.
Guido P.H. Band MSc.
Prof. Jaap F. Hamming
Prof. K. Richard Ridderinkhof

Based on the publication:
Spruit, E. N., Band, G. P., Hamming, J. F., & Ridderinkhof, K. R. (2014). Optimal Training Design for Procedural Motor Skills: a Review and Application to Laparoscopic Surgery. *Psychological Research*, 78(6), 878-891.

Abstract

This literature review covers the choices to consider in training complex procedural, perceptual and motor skills. In particular, we focus on laparoscopic surgery. An overview is provided of important training factors modulating the acquisition, durability, transfer, and efficiency of trained skills. We summarize empirical studies and their theoretical background on the topic of training complex cognitive and motor skills that are pertinent to proficiency in laparoscopic surgery. The overview pertains to surgical simulation training for laparoscopy, but also to training in other demanding procedural and dexterous tasks, such as aviation, managing complex systems and sports. Evidence-based recommendations are provided for facilitating efficiency in laparoscopic motor skill training such as session spacing, adaptive training, task variability, part-task training, mental imagery and deliberate practice.

Introduction

Safe performance in high-stake occupations such as surgery requires extensive and high quality training. In the field of surgery, the use of simulation has been boosted by the introduction of minimally invasive methods such as laparoscopic surgery (LS) (Gallagher & O'Sullivan, 2012). LS requires complex perceptual and motor skills, which need to be acquired efficiently and with high reliability. In search of safe and cost-effective tools for skill acquisition, simulation has been used successfully to train professionals in a wide array of occupations, such as aviation (Salas, Bowers & Rhodenizer, 1998), controlling unmanned systems in the military (Macedonia, 2002; McDermott, Carolan & Wickens, 2012) and in emergency teams in medicine (Shapiro et al., 2004).

Dunlosky, Rawson, Marsh, Nathan and Willingham (2013) have recently provided an extensive review of empirical studies into the efficiency of learning techniques in the domain of knowledge acquisition, comparing techniques like summarization, underlining text, practice testing and distributed practice. Although this review has merit in optimizing the medical curriculum to acquire theoretical knowledge, Dunlosky et al.'s training recommendations cannot be generalized to skill learning. A comparable review of techniques for the acquisition of complex perceptual motor skills would be desirable. However, valuable conclusions about the efficiency of skill acquisition are scattered throughout the literature. The current review aims to summarize evidence-based recommendations for motor skill training. First, the background and training models of LS motor skill training will be described. Next, we relate theoretical concepts of automatization and implicit learning to LS training. In the remaining bulk of the article we highlight several factors and review their validity in the context of LS motor skill training.

A Shifting Paradigm in Medical Training

Surgical skills were traditionally taught in the operating room (OR), in the form of a mentor-apprenticeship model in which senior surgeons provided training to surgical residents (Gallagher & O'Sullivan, 2012). The complexity and difficulty of performing LS forced the surgical community to reconsider this training model, as acquiring these new skills solely in the OR is inefficient and potentially compromises patient safety. LS performance faces several perceptual and motor challenges, as surgeons need to work with spatially extended instruments and an extended camera with a separate video monitor. One of the motor challenges is that the instruments move counter-intuitively as they pivot around an incision in the abdominal wall of the patient (Gallagher, McClure, McGuigan & Sheehy, 2008). This incompatibility between hand motion and outcome is known as the fulcrum effect. Another motor challenge is that LS instruments provide limited degrees of motion freedom (Hiemstra, Kolkman & Jansen, 2008), so surgeons need to think in advance where they place the trocars (incision points) and from what angle their instruments will approach the operation site. In terms of perceptual constraints, several three-dimensional depth cues in vision are lost with the use of a video monitor and tactile feedback is attenuated because surgeons cannot directly feel the surface textures of organs and tissue.

The increase in difficulty of LS as compared to open surgery posed an increased risk for complications. To ensure the quality of patient care, simulation-based training models have been developed so that surgical residents could train in a safe environment. In the past two decades various training models have been designed and tested for construct and predictive validity (Duffy et al., 2005; McDougell et al., 2006; Kolkman, van de Put, Wolterbeek, Trimbos & Jansen, 2008; Seymour et al., 2002; van Sickle et al., 2008).

Government regulations and attention to safety in popular media have pushed further training innovations. In 2008 the inspection of health care in the Netherlands demanded that all teaching hospitals in the Netherlands create a plan of action to improve safety and competence in LS (van der Wal, 2007). In the US, a similar report came out by the Insititute of Medicine in 1999 (Gallagher & O'Sullivan, 2012). Nowadays, training in the US has been formalized in the Fundamentals of Laparoscopic Surgery (FLS) (SAGES, 2013). Similarly, a shift towards the documentation of proficiency in LS skills has taken place in other countries (Aggarwal & Darzi, 2006). It is no longer acceptable that new surgical skills are trained in the OR before residents have shown proficiency in their skills in a safe, simulated training environment. Increases in the complexity of technology and high standards for the quality of health care urge an effective surgical training curriculum, but these increased demands on the curriculum are in conflict with the busy working schedule of surgical residents (Reznick & MacRae, 2006). Knowledge obtained in experimental cognitive and educational psychology does not always find its way to the medical specialists responsible for defining the curriculum, so that even when proficiencies are well established, the most efficient road towards this proficiency is not disclosed (e.g. ASGE training committee, 2012). The goal of the current paper is therefore to review how the acquisition and maintenance of complex LS skills can be facilitated.

Training Models

In motor skill training for LS, various models are available. These can be roughly divided into several categories: physical box trainers, virtual reality (VR) simulators, animal models, cadavers, and patients. A surgical resident usually starts by learning basic LS skills in a physical box trainer and/or VR simulator and practices on animal models and patients at a later stage.

Every training model has its advantages and disadvantages. The box trainer involves characteristic motor tasks with LS tools in an inanimate closed environment and thus has realistic physics, visual and motor requirements and haptic feedback. It is fairly inexpensive, but does not provide summative or immediate feedback like some VR simulators do. Animal, cadaver and patient models provide a more realistic setting, but are also more expensive and have their connected ethical issues. Extensive surveys of different validated simulators and models that are available are provided by Aggarwal, Moorthy and Darzi (2004), Reznick and MacRae (2006), Bashankaev, Baido and Wexner (2011), Schreuder, Oei, Maas, Borleffs, and Schijven (2011) and Gallagher and O'Sullivan (2012).

Several prospective studies have investigated trainee characteristics on the performance and learning curve on LS simulators. For example, it has been demonstrated that Visualization and Spatial Relations (two visuo-spatial ability factors) are predictive of performance on colonoscopy (Luursema, Buzink, Verwey & Jakimowicz, 2010) and laparoscopy VR-simulator training (Luursema, Verwey & Burie, 2012). Also, the age of the trainee has been found to influence the learning rate during LS motor skill training (Risicci, Geiss, Gellman, Pinard & Rosser, 2001; Salkini & Hamilton, 2010). In the current article, we review research findings concerning the variables that apply to the design of the training itself.

Several studies have reviewed the value of LS simulation training (Seymour, 2008; Sturm et al., 2008; Thijssen & Schijven, 2010; Zendejas, Brydges, Hamstra & Cook, 2013). Many studies on laparoscopy training discuss the face validity, concurrent and construct validity of LS simulators, but only a handful focus on predictive validity, most likely due to the excessive time and organization involved in studies that investigate actual improvements in OR performance (transfer) as a result of training (Lynagh, Burton, Sanson-Fisher, 2007; Thijssen & Schijven, 2010).

From a systematic review, Zendejas et al. (2013) concluded that simulation-based training is more effective than video-based instruction alone. Furthermore, for most outcome measures, the use of a physical box trainer leads to better performance than animal models and virtual reality simulators. For example, Avgerinos, Goodell, Waxberg, Cao and Schwaitzberg (2005) found that physical box trainers from the Fundamentals of Laparoscopic Surgery (FLS) teaching and assessment module (SAGES, 2013) discriminated better between different experience levels (juniors, mid-levels, seniors) than did the Minimally Invasive Surgical Trainer - Virtual Reality simulator (MIST-VR). These systematic reviews provide a representative summary of the pooled effects of multiple training factors that are involved in setting up a simulation-based training in LS skills.

In investigating the optimal design of LS motor skill training; we aim to provide recommendations for optimal skill acquisition, retention and to facilitate the ultimate goal of surgical training, transfer to the OR. Bourne and Healy (2012) describe that transfer depends on the similarity of task stimuli and task responses of both the training tasks, as well as the criterion task (real laparoscopic surgery in the OR). If task responses between training and the occupational setting differ too greatly it is possible that performance on the criterion task may even be impaired by training, something for which Gallagher and O'Sullivan (2012) also warn trainers and LS simulation designers. Therefore, optimal design of training during surgical residents' learning period is crucial.

Automatization of skills

Characteristics of the transition period from deliberate, controlled task performance to automatic processing are widely acknowledged in the literature (Logan, 1988; Shiffrin & Schneider, 1977). Initially, fulfillment of a new goal requires full attention to execute a strategy. The benefit of controlled processing at this stage is that task execution is still flexible, but its taxation of cognitive resources makes it slow to execute and hard to combine with other cognitively demanding tasks and this also applies for combining performance and learning. As the task is repeatedly performed following consistent lines (Shiffrin & Schneider, 1977), or as more instances of performance on the task are stored in long-term memory (Logan, 1988), repeating the same task becomes faster, more consistent. And because the need to follow a deliberate strategy decreases, the same actions require less effort with increasing experience. During this process of automatization, capacity for concurrent actions increases, but flexibility in the details of performance is lost. Gallagher et al. (2005) illustrated this automatization process clearly, showing that each additional surgical skill (psychomotor performance, depth and spatial judgments, operative judgment and decision making, comprehending instruction, gaining additional knowledge) requires less attention resources as a laparoscopic surgeon gains more experience in practicing them.

The automatization process provides clues for the opportunities for optimizing training. On the one hand, no shortcuts to mastery or easy alternatives to deliberate practice are readily available. Studies of expertise in fields as diverse as sports, chess, music and handling Morse code consistently show that active involvement in the task for an extended period of time is a prerequisite for attaining high performance levels (Ericsson, Krampe, & Tesch-Römer, 1993). Ericsson et al. showed that, across fields of expertise, on average ten years of intensive experience is required to reach a maximal level of performance. For LS, this implies that once residents make the move to the OR, they have not yet achieved asymptotic performance. On the other hand, there is room for improvement in the efficiency of hours spent in training, especially if the training method is matched to student characteristics such as prior skill level, attention, and endurance, and if the selection of methods is dictated by the ultimate goal regarding safety, speed and flexibility.

Proficiency in surgery involves not only technical motor skills, such as those involved in tying knots and stitching wounds, but also the capacity to rapidly judge unexpected complications and reach adequate decisions on the immediate course of action. The latter capacity is particularly crucial under conditions of fatigue or stress, challenges that are common in the OR. Automatization of technical skills leaves capacity for decision-making under acute stress. One way of accomplishing such automatization in acquiring surgical skill is implicit learning (Masters, Lo, Maxwell, & Patil, 2008). Implicit learning refers to learning rules and their application without being able to explicitly name or describe them (Reber, 1967). Motor skill appears more robust to stress challenges after implicit compared to explicit learning (Mullen, Hardy, & Oldham, 2007). For instance, in surgery trainees, observational learning (vis-à-vis explicit verbal instruction) yields faster stitching performance with fewer hand movements when tested under concurrent-task interference (Masters et al., 2008).

Consistent with the current emphasis on prevention of surgical error (Reznick & Macrea, 2006), errorless learning aims at progressing from easy to more difficult levels of skill, such that errors are avoided altogether during learning. Preventing errors during learning pre-empts the need for generating and testing hypotheses on declarative knowledge (Poolton, Masters & Maxwell, 2005). Discouraging the build-up of such declarative knowledge promotes the build-up of procedural skill. Motor skill is less sensitive to interference from concurrent tasks after errorless (compared to errorful) learning in golf putting (Maxwell, Masters, Kerr & Weedon, 2001) or in hammer nailing in Parkinson patients (Masters, MacMahon, & Pall, 2004). Likewise, errorless learning renders motor skill less sensitive to physical fatigue in learning rugby kicking (Masters, Poolton & Maxwell, 2008; Poolton et al., 2007).

Although initial results in surgical settings are encouraging (Masters et al., 2008), the extent to which the benefits of implicit learning apply also to LS skill learning remains to be further established.

Recommendations for Laparoscopic Motor Skill Training

Setting the Goal of Proficiency

In order to evaluate the efficiency of a training approach, a valid criterion needs to be defined. Although optimal skill acquisition, durability, and transfer to the real occupational setting are ultimate goals, these are not directly assessable following training. Task completion time is a direct indicator of LS proficiency, but it is inadequate as the sole criterion because it is insensitive to safety-compromising errors. Alternatives can be subjective, such as global ratings of video recorded material (Martin et al., 1997); or more objective, such as economy measures of motion (Datta, Chang, Mackay & Darzi, 2002;

Chmarra, Bakker, Grimbergen & Dankelman, 2006), force (Tholey, Desai & Castellanos, 2005; Chmarra, Dankelman, van den Dobbelsteen & Jansen, 2008) and tension in knot tying (Ritter, McClusky, Gallagher & Smith, 2005). Ideally, reliable measures should be obtained for concrete, observable behaviors in differentiable units of behavior of LS performance (Gallagher & O'Sullivan, 2012; Seymour et al., 2002; van Sickle et al., 2008). Although these markers measure different aspects of performance, they are often found to be related. Since the absence of errors is the most crucial factor to patient safety, we view this qualitative measure as the most central marker for LS proficiency.

Specification of the goal of LS training also serves as a didactic purpose. Gallagher and O'Sullivan (2012) advocated proficiency-based training of surgical skills. This approach establishes a benchmark of objectively assessed performance of subject-matter experts on a set of tasks and takes this performance level as a criterion for trainees (Duffy et al., 2005; McDougell et al., 2006; Kolkman et al., 2008). The approach of benchmarking a criterion level of performance of experienced experts has two purposes. First, it helps to establish the construct validity of the surgical simulator as a training device. Second, it creates a specific training goal that can be strived for by surgical residents. Goal setting theory states that performance is elevated when an individual sets specific and difficult goals, as contrasted by a 'just-do-your-best' goal (Locke & Latham, 2002). Using the benchmark expert criterion of performance as a difficult training goal will facilitate the learning process.

Transfer to the real occupational setting ultimately remains one of the most important measures of whether or not a training method has been useful (Mané, Adams & Donchin, 1989), regardless of whether the performance level already improves during training. In spite of the fact that the goal of training is defined in terms of performance, training ought to be designed to facilitate learning, not elevating performance during training per se. This is an important point, because some training techniques (such as massed practice and continual feedback) work to boost performance in the short-term, but will impair long-term retention of skills (Bjork, 1999; Schmidt & Bjork, 1992).

Recommendation 1: Set a proficiency-based benchmark of performance to serve as a learning goal for novice trainees to facilitate skill acquisition.

Adaptive Training and Cognitive Load

Skill acquisition usually benefits from adaptive training regimens. This way, students are not under- or overchallenged during training, as is often the case in a conventional education set-up, where each student is trained under the same conditions regardless of current skill level. Guadagnoli and Lee (2004) recommended that a motor task should initially be predictable, and the difficulty should be low in order to support the development of an initial movement representation by practice. This representation will subsequently serve as the foundation for developing increasingly complex representations of the task, the acquisition of which may involve more delayed feedback and more variable conditions. Ideally, the difficulty of a chosen training task is adapted to the current skill level of a trainee, in such a manner that challenge and skill are well-matched, as this match promotes task engagement (Nakamura & Csikszentmihalyi, 2002). The aim should be to train on a challenging, difficult task that does not result in overwhelming the trainee in terms of cognitive load.

The cognitive load theory provides a framework for attuning the task level to the student's current abilities (Sweller, van Merriënboer & Paas, 1998; van Gog & Paas, 2008). According to this theory, capacity limitations of short-term and working memory determine how much a student learns from a task. Each and every task features some degree of cognitive load, originating from a variety of sources. First, *intrinsic load* stems from the complexity of performing the task itself. Second, *exogenous load* arises from the task environment, such as interruptions, noise or poor instructions. Third, *germane load* refers to the cognitive load of aspects of the task that require skill acquisition (i.e. learning). When cognitive resources are taxed by the intrinsic and exogenous load of a task, very little learning actually takes place. Efficiency can therefore be improved by minimizing exogenous load, and by gradually building up the intrinsic load, as is done by training only parts of a criterion task or scaled-down versions of a task. The use of an adaptive training regimen complies with the recommendations derived from the cognitive load theory, because the challenge level changes as the student progresses.

One adaptive form of criterion task training involves variable prioritization, such that some task elements are held constant while others are set at a lower difficulty. The advantage of this form of training is that it does not isolate the skill elements of a task, a warning commonly expressed by proponents of whole-task approaches (Gopher, Weil & Siegel, 1989). By training only parts of a task, one risks that the elements that are essentially intertwined are not integrated properly when a trainee performs on the criterion task. Gopher et al. (1989) trained participants by shifting the emphasis between the elements that needed to be trained during a whole-task setting. In this way, the emphasis of learning one specific element was maintained, without neglecting the coordinative and integrative demands.

Recommendation 2: Apply adaptive training to allow trainees to progress at their own pace while being optimally challenged in terms of cognitive load.

Part-Task Training

In a part-task regime, training is initially restricted to an incomplete version of the ultimate task (Fadde, 2010; Salden, Paas & van Merriënboer, 2006). Wightman and Lintern (1985) distinguished between ways in which a complex task can be broken into its components, such as segmentation and fractionation.

Segmentation implies that the specific actions in time and space are separated and put into parts of a task. In LS, this would refer to different sub-steps of a procedure (segmentation of temporal action sequences) and handling the left and right hand instruments (spatial segmentation). Eventually, temporal segments need to be recombined to train coordination and integration. Different ways of chaining previously trained motor actions into larger motor chunks have been described (Salden et al., 2006), but will not be discussed here in detail.

Fractionation, in contrast, refers to deconstruction of a task into individual elements that eventually need to be performed in combination. For example, Mané et al. (1989) studied part-task training of virtual ship navigation and a defense task, which later needed to be combined. This reduced the required training time and did not reduce the transfer of performance from fractionated tasks to the whole task relative to a whole-task training group. In this way, the part-task emphasis of learning one specific element is still maintained, but subjects also learned how to integrate all the other elements of the task. It is important to emphasize that the benefits of part-task training as described by Mané et al. (1989) cannot be generalized to all forms of fractionating a complex task (Wickens, Hutchins, Carolan, & Cumming, 2013). If task parts need to be coordinated in the criterion task, pure part-task training has negative transfer effects, presumably because the control requirements of integrating task components also need to be trained (Wightman & Lintern, 1985), whereas task parts that are only sequentially combined in the criterion task show a clearer benefit of part-task training (Peck & Detweiler, 2000).

Fractionation has been applied implicitly in the design of LS motor skill tasks, expressed in the various tasks that exist on surgical simulators (such as camera navigation, tissue handling, clipping, cutting, bi-manual cooperation, intracorporeal suturing, etc.). In fact, the most commonly used simulators in LS training are themselves a form of fractionation as they primarily train motor skills, while often neglecting perceptual, social, technical and decision-making skills that are important elements of expertise in real LS. This is noteworthy because there are few studies addressing the predictive validity of surgical simulators (Lynagh et al., 2007; Thijssen & Schijven, 2010). Part-task training by means of segmentation has also been applied implicitly on VR-models in which sub-steps of an entire LS procedure can be performed, but it has yet to be implemented in the context of learning a complex motor skill (such as intracorporeal suturing). The potential advantages of part-task training have yet to be tested in empirical studies on LS training.

Recommendation 3: Use part-task training for tasks that can be segmented, but not for fractionation of tasks that have many interacting elements.

Dual-task training

Training LS skills is intended to reduce the burden of a task by automatized performance, so that the surgeon retains capacity for additional cognitive demands, such as interacting with colleagues during an actual LS operation. Quantification of the diminishing burden is therefore valuable, to evaluate both the effectiveness of an approach and the trainee's development. Subjective measures of mental workload can take the form of questionnaires and interviews, but these rely on the validity of self-reports. More reliable measures of automatization are psychophysiological measures indexing effort and arousal exerted on a LS task (Carswell, Clarke, & Seales, 2005) such as ocular and cardiac measures, but also thermal imaging technology (Pluyter, Rutkowski, Jakimowicz & Saunders, 2012).

Alternatively, a secondary task can be introduced to either probe the availability of resources, or to test the robustness of the automatized task. For example, Goodell, Cao and Schwaitzberg (2006) presented a secondary arithmetic task while novices trained on a MIST-VR simulator. This additional demand interfered with completion time, but not with the economy of movement or error rate. Hsu, Man, Gizicki, Feldman and Fried (2008) let LS novices and experts combine a LS peg transfer task as the primary and math problems as the secondary task. Novice trainees managed to maintain performance on the primary task, but did so at the expense of performance on the secondary task, whereas the experts where able to maintain performance on both tasks. This finding can be explained by a difference in the level of automatization of the LS skills, leaving experts spare attention resources to cope with the cognitive demands of a secondary task. These findings show how demonstrated proficiency on a single LS task may not be sufficient to decide whether a trainee is ready to make the transition to the more demanding OR.

An additional advantage of introducing a secondary task in training is that it trains the task-coordinative skills that are indispensable in the OR. Mann (2011) referred to this as situated learning, which is advocated based on the assumption that any form of learning is always linked to the context in which it occurs. Teague, Gittelman and Park (1994) reported that transfer from training to the criterion task is improved when the context environment of learning and transfer are comparable. For context-dependent performance, it is better to train in a similar learning environment, whereas for performance that needs to transfer to a broader range of contexts, it is better to train in a context-independent learning environment. LS motor skill training remains a form of task fractionation of actual surgical performance in the OR. Adding other tasks during training and assessment can help to integrate skills that were initially trained in isolation in order to facilitate transfer to the real surgical context.

Pluyter, Buzink, Rutkowski and Jakimowicz (2010) found that medical trainees' performance significantly decreased on a LS task with a more realistic high cognitive load training design. Participants had a significantly lower task score and higher error rate on a clipping and cutting VR task when confronted with background radio noise and a poor assistant navigating the laparoscope. Trainees also reported that they experienced more stress during the task under the more realistic training conditions as compared to the artificial low cognitive load conditions. These findings suggest that increasing the cognitive load during LS motor skill assessment can be a valuable strategy for improving transfer of acquired skills to the OR.

Recommendation 4: Use dual-task training to assess the degree of automaticity of skills.

Recommendation 5: Adopt context-specific elements from the criterion task into the training to support positive transfer of skills to the occupational setting.

Task Variability

Automating performance is associated with a reducing involvement of deliberate control, and this increases the rigidity of task execution. To retain the flexibility that is required to adapt to unanticipated task demands, one possible solution is to train with task variability. Schmidt and Bjork (1992) argued that performance in the short-term tends to be enhanced when task parameters (such as distance and speed) within a motor task are held constant, but that long-term retention of skill is facilitated when these parameters are variable. This is a relevant finding, since many simulated tasks for LS training are performed under fairly constant conditions, whereas real surgery in OR will provide plenty of variations in terms of anatomy (type of disease, sex differences, amount of fat, scar tissue formation, anomalies, etc.). This finding again highlights the more general notion that the focus during training should be on skill acquisition and long-term retention, not on enhancing performance during training.

An interesting research question is whether introducing more variability in LS training tasks will yield more learning benefits. In the literature on task variability during motor skill training, it is commonly observed that changing the training task frequently improves learning outcomes (Proteau, Blandin, Alain & Dorion, 1994). This is called the contextual interference effect, and refers to the observation that variable conditions of practice during a block of time impairs performance during the skill acquisition phase, but improves performance on a delayed retention test as compared with practicing only on the exact same task for repeated trials during a block of time. One explanation for this is that on a more variable schedule of practice, trainees have to hold multiple tasks in working memory which allows for elaboration and comparing and contrasting them with each other (Boutin & Blandin, 2010). Another explanation is that trainees have to reconstruct their motor action plan for each specific task more often due to the high rate of task switching during training, which results in enhanced learning.

This effect has been found to improve verbal skill learning, as well as motor skills. For motor skills, contextual interference can be increased by task switching between task that are controlled by different motor programs, rather than just creating small variations of a task that relies on similar motor programs (Magill & Hall, 1990).

Jordan, Gallagher, McGuigan, and McClure (2000) tested the role of task variability in LS skill acquisition. They showed that trainees automatized faster to the fulcrum effect when they were presented with variable conditions of the video screen of the LS simulator they practiced on. The results showed that when randomly alternating a regular and mirrored version of the screen during training, trainees habituated to the inverse moving instruments faster, as compared with trainees who only had the regular or mirrored version of the screen. This is a clear example of the benefits of introducing variability of task parameters. In future studies, the effects of greater level of contextual interference can be tested to verify the validity of this effect in LS motor skill training.

Recommendation 6: Introduce task variability to enhance skill acquisition and long-term retention and to optimize flexibility of trained skills.

Spacing versus Massed Training

To obtain the best training results given a time investment, an important choice to make is whether training can best be scheduled as long blocks of training, or rather as multiple short training sessions. Dunlosky et al. (2013) reviewed the evidence in knowledge acquisition, showing that spaced sessions are more effective than massed learning, and this also seems to apply to motor skill acquisition (Shea, Hai, Black & Park, 2000). One of the factors pertaining to this effect is the role of mental fatigue in learning (Van der Linden, 2011), which is known to affect psychomotor and cognitive skills in LS tasks (Kahol et al., 2008). The length, planning and distribution of training sessions thus affect the availability of cognitive resources during training, resulting in a depletion of them in sessions that are too long in duration. The detrimental effect of fatigue can be reduced by varying the task load and also by limiting session length.

Mental fatigue is not the only argument against massed training. This is indicated by an additional effect of lag duration on long-term retention, that is, more time transpiring between separate practice sessions supports memory, at least for explicit knowledge acquisition (Dunlosky et al., 2013). One way to explain the advantages of spacing is by the power law of forgetting, in which the accuracy of memory deteriorates as learned material decays over time (Wixted & Carpenter, 2007). On spaced practice sessions, a trainee has to exert more effort than in a massed schedule to attain criterion performance. This forces a trainee to invest more cognitive resources into learning, with a positive effect on long-term retention.

Another explanation for the spacing benefit is that trainees tend to overestimate how well the material has been learned if it is easier to retrieve studied material after a short time interval as compared to a longer time interval (Dunlosky et al., 2013). Bjork (1999) argued that this misplaced confidence is also a risk in fixed and predictable conditions of training. These training conditions improve performance during training, but increase the risk of inaccurate appraisal of how well the material has been learned. The concept of appraisal of learned material is central in training, especially when practice is self-directed. Son (2004) found that students spaced and massed their study material on the basis of how difficult they judged to-be-learned items to be, and this is in turn affected retention. Thus, shortening training sessions and spacing practice will lead to a more accurate appraisal of skill level.

Additionally, skill acquisition does not only occur during the training session itself (online periods), but also in the time period following training (offline periods; Hallgató et al., 2012). Korman et al. (2007) describe that trainees have two gains in performance, one within-session during training and one between-session in a process called consolidation that occurs after training. Typically, the gains of consolidation accrue in the window of a couple hours directly following training, as well as during overnight sleep. During the first time window, retrograde behavioral interference can occur, which means that motor patterns that are practiced later in training will impair the consolidation of the motor skills that were practiced initially in training (Brashers-Krug, Shadmehr & Bizzi, 1996). This effect can be partially mitigated by a nap in between training of different motor patterns (Korman et al., 2007), indicating that sleep plays a very important part in the degree of consolidation in both the immediate time window following training, as well as the overnight period.

Loehr and Schwartz (2003) stated that recovery of one modality can take place during active participation in a task that requires a different modality. For example, engaging in physical activities in leisure time promotes recovery from work that is high in cognitive demands (Rook & Zijlstra, 2006). Spacing different kinds of training and work activities can help to maintain optimal engagement in both work and training. Obviously, many considerations have to be taken into account when planning the working schedule of a surgeon, but where possible, spacing different types of work activities should certainly be one of them.

In a meta-analysis, Donovan and Radosevich (1999) differentiated the spacing effect by task type and concluded that spaced practice has larger benefits for simple motor tasks than for complex tasks, including tasks with a strong motor control load. Nonetheless, benefits also apply to the medical field. Moulton et al. (2006) found that spaced practice was superior to massed practice in skill acquisition and transfer for microvascular anastomosis, and the value of spacing training has been replicated successfully for learning LS skills using the MIST-VR (Gallagher, Jordan-Black & O’Sullivan, 2012). Our own research group has obtained similar results in a study of training intra-corporeal suturing with a LS box trainer model (Spruit, Band & Hamming, in preparation). These findings have implications for the current training curriculum in medical centers, as most of these present their LS courses on a massed schedule. Spacing in between training sessions should ideally be around 15-20% of the total time to the moment of required transfer to the occupational setting for short time intervals (weeks), although this ratio becomes lower (around 8%) for longer retention intervals (years) (McDaniel, 2012). Although massing training is more convenient in terms of logistics, it may be less cost-effective for a trainee's skill acquisition and retention. When taking into account the differences in learning on a massed versus spaced training program, the latter will prove to be more efficient.

Recommendation 7: Make use of spaced training sessions to facilitate skill acquisition and long-term retention and to avoid inaccurate appraisal of perceived skill level.

Mental imagery

Training on complex skills need not be restricted to physical training sessions to achieve progress. In sports psychology it is widely acknowledged that training by mental imagery of actions without actually enacting them has its own merits (Cumming & Hall, 2002; Brouziyne & Molinaro, 2005). Cumming and Hall showed that athletes activated comparable brain areas during imagery and physical training, suggesting activation of similar cognitive processes. Mental imagery can have different cognitive and motivational functions; such as imaging overall plans of action, mental rehearsal of motor skills, focusing on the sensations experienced during a given task, using imagery to enhance feelings of confidence and focusing on the imaging of successful achievement of a desired goal or performance (Hall, Mack, Paivio and Hausenblas, 1998). All these functions are relevant for LS skills training. Moreover, mental imagery is independent of constraints to the availability of resources and can be performed anytime, anywhere.

There are different ways of performing mental imagery (Immenroth et al., 2007). Typically, a trainee starts by observing an example, a model that shows how to perform a motor skill successfully. After this, the trainee can later try to recall an image of the model performing the skill and create a mental image of him or herself executing the task; the more vivid, the better. This mental image can either be associated or dissociated, meaning that it is either viewed from an internal first-person perspective or from an external third-person perspective.

In addition, the focus of the trainee's attention can be internal or external in motor imagery; a distinction that is also made for motor execution. Wulf, Shea and Lewthwaite (2010) recommended an external focus of attention on movement effects during training, which is in line with the advice to use external mental imagery for novices learning a new LS motor skill (Whiting & den Brinker, 1981). By focusing on the outcome of an action, the action-effect coupling is thought to be trained. Cumming and Ste-Marie (2001) proposed that the relative value of an internal versus external mental imagery perspective may depend on task-specific characteristics and the level of experience of the performer. Motor skills aimed at the appearance of the movements (such as gymnastics and figure skating) may benefit more from external imagery, which mostly engages the visual modality. In contrast, motor skills that require ongoing processing of incoming information (such as canoe slalom and car racing) is facilitated more by internal imagery, which takes into account the visceral and kinesthetic sensations.

An external perspective may be better suited for a beginner just learning the general motions of certain motor actions, whereas an internal perspective is more appropriate at a more advanced stage in which a performer is homing in on the details of a motor skill (Whiting & den Brinker, 1981). The evidence regarding these ideas is inconclusive and requires domain-specific research as findings in one type of task do not necessarily generalize to another type of task. However, the internal-external perspective distinction could have implications for setting up training in surgical skills as real LS in the OR would certainly qualify as a skill that requires accurate perceptual information (correct identification of anatomical landmarks and awareness of where one is operating). On the other hand, novices just learning a new basic surgical skill in the lab would be better suited to use more external mental imagery, because this type of imagery will facilitate a beginners training process better than internal imagery.

Sanders, Sadoski, Bramson, Wiprud and van Walsum (2004) set up a study examining the use of mental imagery rehearsal in the practice of suturing. The authors found that participants who used mental imagery instead of an additional physical practice session performed equally well as the group with more physical practice. This indicated that mental imagery practice is a good tool for increasing cost-effectiveness. These findings were replicated for venipuncture (Sanders et al., 2007), lumbar puncture (Bramson et al., 2011) and showed that mental imagery was superior to text-book study (Sanders et al., 2008).

Immenroth et al. (2007) examined the effect of mental training on LS cholecystectomy in a porcine model. Results indicated a higher increase in performance for a mental-training than for the control group on a task-specific procedural checklist, capturing cognitive elements, but not on global rating scales from OSATS (Objective Structured Assessment of Technical Skills), capturing motor skills. Arora et al. (2011) studied mental practice on LS cholecystectomy in a VR simulation. Results indicated superior scores for the mental training group on both OSATS and a validated Mental Imagery Questionnaire. Taken together, these findings show that mental imagery practice can be embedded in a training curriculum for LS to improve efficiency and cost-effectiveness. Additionally, Cumming and Hall (2002) argued that mental imagery may qualify as a form of deliberate practice that enhances skill acquisition.

Recommendation 8: Train the use of mental imagery to effectively reduce training costs.

Deliberate Practice: From Proficient to Expert

Gallagher and O'Sullivan (2012) derived their proficiency-based training recommendations from a model of skill development that distinguishes five stages: Novice, Advanced Beginner, Competent, Proficient, and Expert (Dreyfus, Dreyfus, & Athanasiou, 1986). Although it is unrealistic to expect a surgical trainee to reach the Expert stage during their period of training, it is vital to examine which qualities, values and conditions inevitably lead to this stage and to incorporate these into the training curriculum. Ericsson (2006) stated that extended practice on a skill will lead to proficient and functional performance on a task, but will not invariably lead to expert levels of performance. For example, Plant, Ericsson, Hill, and Asberg (2005) found that the amount of study time was not predictive of academic performance, but the quality of study time and prior performance were better indicators.

Traditionally, the difference between proficient and expert levels of skill has been attributed to factors such as innate abilities, cognitive capacity and talent (Coyle, 2009). Ericsson (2006) posed a counter-argument, stating that this difference is explained by the way experts continue to practice as their experience accumulates. Whereas most trainees start to perform their tasks by means of automatization as they reach proficient skill levels, the practitioners that reach the peaks of expertise make sure to keep their cognitive control engaged in honing their skill level. This form of practice is called deliberate practice (DP) and consists of consciously designing practice that is aimed at improving a highly specific aspect of performance, usually a sub-skill that a trainee is dissatisfied with and would lead to the greatest benefit if focused upon during practice.

For medical education, McGaghie, Issenberg, Cohen, Barsuk, and Wayne (2011) reviewed studies that compared simulation-based training using a DP protocol with traditional clinical medical education. They found that simulation-based training combined with DP yields superior skill acquisition. Crochet et al. (2011) tested the value of DP in a VR-simulator training of LS cholecystectomy. A DP group received individualized feedback on their strong and weak points and performed training drills aimed at improving that specific aspect of their performance, whereas the control group watched surgical tutorials. Both groups improved in performance on the VR-simulator and their skills transferred to LS in a porcine model. Although the control group was significantly faster, qualitative rating scales and procedure checklist indicated superior performance for the DP group. The authors explain these findings in terms of a sacrifice of dexterity (time taken) for quality of performing the procedure (Crochet et al., 2011).

Although this explanation is certainly plausible, it is questionable whether research on top level expertise and DP lends itself to a short intervention study. The road to expertise is a long one, with Ericsson (2006) indicating around 10,000 hours of DP. It is unlikely that an effect that adds up gradually over years of practice can be validated in a study comprising of several weeks of motor skill training. This learning period applies mostly to the learning curve from novice to competent (or perhaps proficient), but not to expert. Much research on expertise and DP is oriented on qualitative inquiry into what makes the difference between proficiency and top level expertise and the principles derived from this are hard to validate in quantitative studies simply due to time constraints. None-the-less, we suggest that it is essential to incorporate the basic principles of DP to facilitate learning during the phase from proficient to expert.

Some of the elements of DP are already described in this paper (such as well-defined proficiency goals, matching task difficulty level to the current skill of the trainee and reliable assessment), others include informative feedback from simulators and trainers, monitoring performance, error correction and high trainee motivation and concentration (McGaghie et al., 2011). The core value of DP however lies in the long-term commitment to an ongoing process of learning that doesn't just rely on an external benchmark, but on an internal drive to continually refine surgical skills. Proficiency-based training alone will set a context for trainees to practice their skills up to a certain level (the benchmark) and automatize their skills up to that point. Endorsing principles of DP will create trainees that focus on career-long improvement and reduce deterioration of skill after training has been completed.

Attention to different stages of skill acquisition is important, as it seems that many of the cognitive load-reducing learning strategies that are effective for a novice can become detrimental for more skilled trainees; this has been referred to as the expertise reversal effect (van Gog, Ericsson, Rikers & Paas, 2005). Care should be taken to individualize training towards the specific needs of particular trainee, but also to deliver the instructions of training according to a trainee's current skill level.

Recommendation 9: Incorporate principles of deliberate practice to foster career-long improvement of skills in trainees.

Generalizability to Other Fields of Motor Skill Training

Recommendations in this literature review were primarily geared towards LS motor skill training. Despite this narrow focus, many of the training principles can be extended to related fields that involve the training of complex perceptual and motor skills. For a learning strategy such as spacing, the advantages are very robust; not only for knowledge, but also for skill acquisition. Other training variables require more nuance, such as part-task training, the effectiveness of which differs greatly for tasks with large time-sharing concurrent skill components and tasks that are more sequential (Wickens et al., 2013). The nature of the motor task is important to consider, since tasks that involve responding intuitively to sensory stimuli and online error correction (such as boxing) rely on different functions than tasks that involve extensive motor sequence planning (such as gymnastics) (Purves et al., 2012). Research on training variables in each particular discipline is vital in proving the effectiveness of the learning strategies in differing contexts of motor skill.

CHAPTER 04

FEEDBACK IN THE LEARNING PROCESS OF LAPAROSCOPIC SKILLS

Edward N. Spruit MSc.
Dr. Guido P.H. Band
Prof. Jaap F. Hamming

Based on the publication:
Spruit E.N., Band G.P.H. & Hamming J.F. (2013), Feedback bij het aanleren van laparoscopie, *Nederlands Tijdschrift voor Geneeskunde* 157(23): A6354

Abstract

A recent Danish study showed that instructor feedback significantly reduced the duration of training time needed for acquiring laparoscopic skills. While there is a clear advantage to trainees reaching a predetermined expert level of performance more rapidly, this does not necessarily imply that the skills were also acquired more efficiently. Experiencing continual feedback while undergoing a training task could reduce the level of difficulty in performing it; the presence of an instructor can also heighten emotional tension. Both of these factors can impair the learning process. For this reason, we recommend self-directed feedback during training on complex laparoscopic skills.

Feedback in the learning process of laparoscopic skills

The value of training laparoscopic skills in a safe and controlled environment has been confirmed in several studies (Seymour, 2008; van Sickle et al., 2008). An important question when designing laparoscopic training programs is what kind of feedback is optimal for trainees. The effect of feedback depends on the timing of it and can be different for each type of task (Healy & Bourne, 2012). Delayed feedback improves feedback for mental tasks, while delaying feedback impairs performance on motor tasks.

Feedback in laparoscopy training is most effective when it is presented to the trainee shortly after making an error (proximal feedback), but not in a summary form at the end of the task (summative feedback) (Gallagher & O'Sullivan, 2012). Most virtual reality simulators which are developed for laparoscopy training present summative feedback to trainees, without any specified directions on how a mistake was caused or suggestions of any alternative better strategies. These kinds of suggestions can be given in the form of proximal feedback by an experienced instructor.

Feedback by an instructor

A recent Danish study shows that trainees require less training time to reach a predefined proficiency level when an instuctor provides them with feedback (Strandbygaard et al., 2013). The authors showed this finding using a virtual reality simulator for laparoscopic salpingectomy. However, the performance score of trainees that did not receive any feedback (control group) was higher than in the group that did receive feedback. The authors attribute this result to the fact that the total amount of training needed to reach the predefined proficiency level was also higher in the control group.

This finding is in line with the idea that performance score are not the most important measure of an effective training (Mané, Adams & Donchin, 1989). Feedback by an instructor may help a trainee to reach a predefined proficiency level faster, but this does not mean a trainee also learned the tasks better necessarily. In other words: a trainee is usually quite capable to perform a complex laparoscopic task when an instructor is talking him/her through the procedure step by step, but the trainee is often not able to reproduce the same set of actions when this line of help disappears.

Because learning requires cognitive resources (next to execution/practice), the saying 'learning by doing' isn't always as viable. When the scope and complexity of a task are large, the execution of a task can be a large enough tax on cognitive resources to an extent that there will be no capacity left for learning (Healy & Bourne, 2012; Paas, Renkl & Sweller, 2003).

The Danish research Group states a trainee is better off without feedback while performing simple laparoscopic tasks, while feedback on complex tasks is desirable (Strandbygaard et al. 2013). The authors state the feedback ought to be self-directed, in the sense that a trainee can suggest themselves whenever they would like to receive feedback. In this way, a trainee can determine their own pace, which is a hallmark for the success of adaptive training towards a predefined proficiency level (Gallagher & O'Sullivan, 2012; Mané, Adams & Donchin, 1989).

Pros and Cons

Feedback by an experienced instructor has both advantage and disadvantages. An advantage is that an instructor can see directly what trainee is doing incorrectly and can provide specified and individual feedback. Also, interpersonal factors have an influence. A committed instructor knows the value of the learning process and cares that a trainee will be fully engaged in training. An instructor can also affirm a trainee's progress and help to motivate whenever a trainee becomes frustrated or discouraged.

However, it is undesirable that an instructor continues to talk a trainee through the procedure, providing a constant safety net. Correct instruction and continuous feedback can contribute to a trainee's process in the sense that they do not learn any wrong habits (Healy & Bourne, 2012), but they shouldn't cause a trainee to become dependent and unable to perform a task on their own.

The presence of an instructor has an influence on the arousal levels of a trainee. It is a well-documented finding that the execution of well learned tasks improves and the execution of poorly learned tasks deteriorates in the presence of others (Zajonc, 1965). This can be explained by the fact that a more difficult task creates a higher cognitive and emotional demand on a novice trainee than an easy task. The presence of other people creates additional arousal, which cause performance to decline. This is why most people prefer not to have any other people watching them while they are still learning a new skill. When a skill has been mastered, the observation of others has a positive effect, because it increases the engagement of the trainee. This social phenomenon has implication for designing laparoscopy training in a skillslab, but also for the traditional Halsted-model while teaching residents in the operating room (Gallagher & O'Sullivan, 2012), where even more people are usually present. This additional arousal may impair the learning process of a trainee and the safety of a patient.

Finally, the presence of an instructor costs time and money. Because trainees commonly make similar mistakes, it can often occur that instructors are finding themselves repeating similar feedback to multiple trainees. This can be solved by more efficient alternatives. Our research group tries to cover most common newbie mistakes in standardized instructional videos, which are presented to trainees multiple times during training.

Recommendation

Our advice is to apply self-directed feedback by instructors while teaching trainees challenging laparoscopic skills with a predefined proficiency level, so trainees can determine their desire for feedback themselves. Besides experienced surgeons, training alumni, lab assistants and even medical students can be recruited as instructors, granted that they possess proficient laparoscopic skills.

CHAPTER 05

IMPROVING TRAINING OF LAPAROSCOPIC TISSUE MANIPULATION SKILLS USING VARIOUS VISUAL FORCE FEEDBACK METHODS

Daan Smit MSc.*
Edward N. Spruit MSc.*
Dr. Jenny Dankelman
Dr. Gabrielle Tuijthof
Prof. Jaap F. Hamming
Dr. Tim Horeman
*: Both authors contributed equally to this chapter

Based on the publication:
Smit, D., Spruit, E. N., Dankelman, J., Tuijthof, G., Hamming, J. F., & Horeman, T. (2017). Improving training of laparoscopic tissue manipulation skills using various visual force feedback types. *Surgical Endoscopy*, 31(1), 299-308.

Abstract

Background: Visual force feedback allows trainees to learn laparoscopic tissue manipulation skills. The aim of this experimental study was to find the most efficient visual force feedback method to acquire these skills. Retention and transfer validity to an untrained task were assessed.

Methods: Medical students without prior experience in laparoscopy were randomized in three groups: Constant Force Feedback (CFF) (N=17), Bandwidth Force Feedback (BFF) (N=16) and Fade-in Force Feedback (FFF) (N=18). All participants performed a pre-test, training, post-test and follow-up test. The study involved two dissimilar tissue manipulation tasks, one for training and one to assess transferability. Participants performed six trials of the training task. A force platform was used to record several force parameters.

Results: A paired sample T-test showed overall lower force parameter outcomes in the post-test compared to the pre-test ($p < .001$). A week later the force parameters outcomes were still significantly lower than found in the pre-test ($p < .005$). Participants also performed the transfer task in the post- ($p < .02$) and follow-up ($p < .05$) test with lower force parameters outcomes compared to the pre-test. A one-way MANOVA indicated that in the post-test the CFF group applied 50% less Mean Absolute NonZero Force ($p = .005$) than the BFF group.

Conclusion: All visual force feedback groups showed to be effective in decreasing tissue manipulation force as no major differences were found between groups in the post and follow up trials. The Bandwidth Force Feedback method is preferred for it respects individual progress and minimizes distraction.

Introduction

Although laparoscopic surgery brings many advantages for patients (smaller scars, shorter hospitalization), the disadvantages are to the extent of surgeons as task complexity increases. Laparoscopic surgery requires more of the capabilities of surgeons compared to open surgery (Blavier, Gaudissart, Cadière & Nyssen, 2006; Spruit, Band, Hamming & Ridderinkhof, 2014, chapter 3). Tactile feedback is degraded as a consequence of instrument friction (Ahlberg et al., 2007; Sinitsky, Fernando & Berlingieri, 2012) and between instrument and trocar (van den Dobbelsteen, Schooleman & Dankelman, 2007). Psychomotor challenges, such as counter intuitive movement (fulcrum effect) of the instruments (Ahlberg et al., 2007; Sinitsky, Fernando & Berlingieri, 2012) and limited degrees of motion freedom (Blavier et al., 2006; Horeman, 2014), contribute to the increased complexity of the operating technique as well. Even though safe tissue handling is an important topic in the training of surgical skills, it is difficult to assess. New training and assessment methods were developed and validated (Horeman, Dankelman, Jansen & van den Dobbelsteen, 2014) and used to provide a more objective measure for the "instrument handling" and "tissue manipulation" grading sections as used in the OSATS scoring form (Martin et al., 1997).

Previous research shows that visual force feedback contributes to safe tissue manipulation (Horeman, Blikkendaal et al., 2014; Horeman, van Delft et al., 2014). However, providing frequent or continuous presentation of visual feedback does not consistently contribute to the learning process, in some cases it may even hinder skill acquisition (Buchanan & Wang, 2012; Lam, DeRue, Karam & Hollenbeck, 2011; Magill, 2007; Patterson, Carter & Hansen, 2013). High frequency feedback guides the trainee to correct movement (Wulf, 2007), but overexposure can create feedback dependency (guidance effect) (Buchanan & Wang, 2012; Lam, DeRue, Karam & Hollenbeck, 2011; Magill, 2007; Patterson, Carter & Hansen, 2013). This can lead to fluctuation in performance because the trainee is constantly correcting small, insignificant errors (Wulf, 2007).

An obvious solution to overcome the guidance effect is to omit continuous feedback (Buchanan & Wang, 2012). This will strengthen the intrinsic ability to discriminate between skill effective and ineffective behaviour and decreases dependency on feedback (Buchanan & Wang, 2012; Patterson et al., 2013). In this study, we aim to apply this theory by evaluating two different methods of lower frequency feedback for laparoscopic skills training in box simulators.

Fade-in feedback

In the literature, a number of options are suggested. One of those options to solve the guidance effect is fade-in feedback (Magill, 2007; Sigrist, Rauter, Riener & Wolf, 2013).

Feedback can possibly be overwhelming for the performer at the start of training (Spruit et al., 2014, chapter 3) if it exceeds the attention capacity at the beginning of the acquisition process. The trainee therefore should only be presented with feedback when the surgical task demands less conscious attention of the performer (when the task has become automated).

Bandwidth feedback

Another proposed option to undermine the guidance effect is bandwidth feedback. In this setting, the trainee will only be presented with feedback when his or her performance exceeds a certain threshold (Ribeiro, Sole, Abbott & Milosavljevic, 2011) and thus respects individual progress (Sigrist et al., 2013). Of major importance is establishing the threshold, the tolerable amount of error before confronting the trainee with feedback. Adverse thresholds will result in over- (i.e. results in unstable set of execution skills) or underexposure (i.e. results in skill execution which contains errors) to augmented feedback and may lead to suboptimal performance (Sigrist et al., 2013).

The aim of the current study is to determine the most efficient dosage of visual force feedback using Constant Force Feedback, Fade-in Force Feedback and Bandwidth Force Feedback.

Method

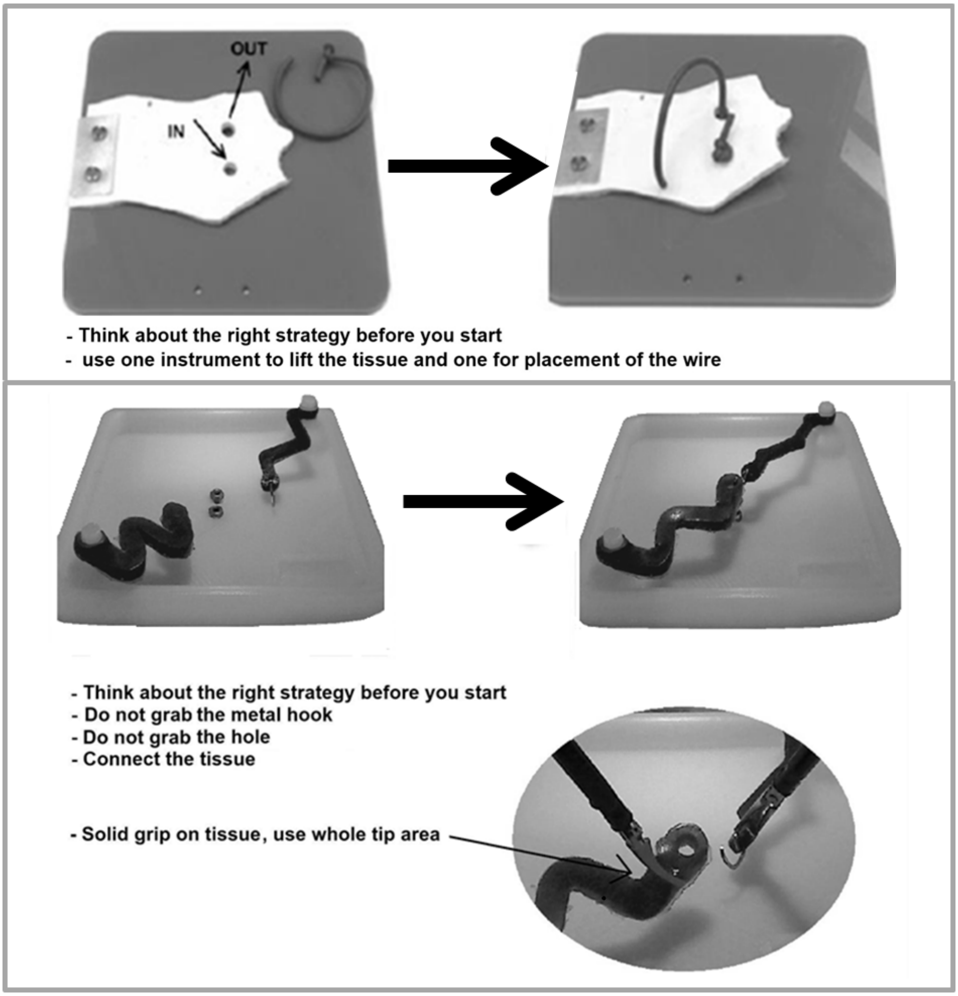
Participants

Medical students without prior experience in laparoscopy training were recruited for the study. The study included 51 participants (30 women; mean age 19.69, range 17-30) of which 1 participant did not turn up for the follow-up test. Participants were assigned semi-randomly to one of the three groups, based on their availability. Furthermore, it was unknown for the participants that each timeslot available for training had a predefined group protocol assigned to it. The Constant Force Feedback (CFF) group consisted of 17 participants (6 men, 11 women; mean age 20.12, range 18-24), the Bandwidth Force Feedback (BFF) group consisted of 16 participants (10 women; mean age 19.63, range 17-30) and the Fade-in Force Feedback (FFF) group consisted of 18 participants (9 women; mean age 19.33, range 17-28).

Test setup

The ForMoST hybrid trainer is equipped with the TrEndo tracking system, the ForceTRAP force tracking system and an USB camera for the visualization of the task on the computer screen (Horeman, Rodrigues, van den Dobbelsteen, Jansen & Dankelman, 2012). The

ForMoST system measures all instrument movement and forces exerted on the training task.



- Think about the right strategy before you start
- use one instrument to lift the tissue and one for placement of the wire

- Think about the right strategy before you start
- Do not grab the metal hook
- Do not grab the hole
- Connect the tissue

- Solid grip on tissue, use whole tip area

Figure 1. The instructions for Task 1 (top) and Task 2 (bottom) as they were presented on the display of the hybrid trainer.

Tasks

To assess the surgical skills required for proper tissue handling, two tasks validated for force parameters were used (Horeman, Dankelman et al., 2014; Horeman, van Delft et al., 2014) which make use of elastic elements that mimic properties of real tissue. Bimanual cooperation of the instruments is essential to complete both tasks.

Task 1: The objective of the task was to guide the wire completely through the two holes of the patch, using a predefined route (Figure 1). The task is designed to force participants to work bimanually with both instruments. If Task 1 is performed correctly, the applied force is negligible.

Task 2: In order to complete Task 2 successfully, connection of the silicone strips should be accomplished with insignificant exerted force. Different from the original task as described in our previous work, the two silicon strips differed in shape and stiffness to make the participants aware that tissues in the human body differ as well. Figure 1 shows the instructions provided to the participants before the pre-test measurement was started.

Study design

Participants performed the two different training tasks inside the ForMoST hybrid trainer. Task 1 was used in the pre-test, post-test and follow-up test (Figure 2). Task 2 was used in the pre-test, training, post-test and follow-up test. Task 1 was used to observe if the force feedback training with Task 2 generated transfer to Task 1 indicated by a decrease in force parameter outcomes values. The study consisted of two meetings. The duration of the first meeting was 90 minutes and the second meeting, scheduled one week later, had a duration of 15 minutes.

The training consisted of 6 trials of 5 minutes each. Participants received real time visual force feedback during training according to the force feedback group assigned to. The CFF group received continuous feedback about their applied force. Participants in the BFF group were only presented with visual force feedback when their applied force exceeded the threshold of 5.3 Newton. The threshold was based on a previous study that defined the critical force level that cause tissue damage (Rodrigues, Horeman, Dankelman, van den Dobbelsteen & Jansen, 2012).

Once the visual force feedback was presented, it lingered for 10 seconds to give the participants the opportunity to notice the feedback and to correct their actions accordingly. The presented force feedback then disappeared again, but only if the exerted force was decreased below the threshold of 5.3 Newton. The FFF group were not exposed to force feedback in their first training trial. In the second training trial participants were presented with force feedback solely in the first minute. The time force feedback was presented gradually increased every trail by a minute. In the last training trial participants of the FFF group were continuously presented with force feedback.

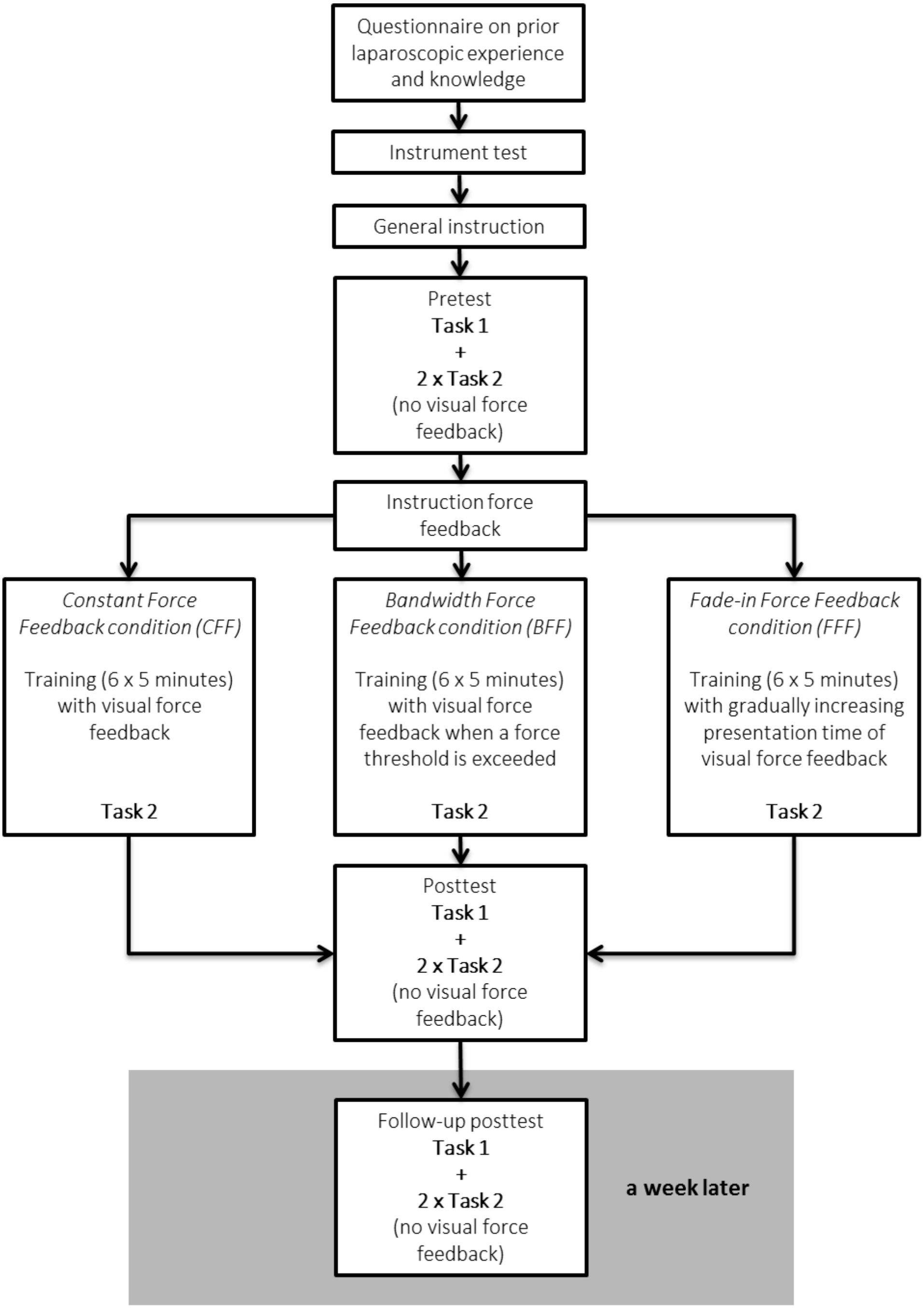


Figure 2. Schematic view of the study design.

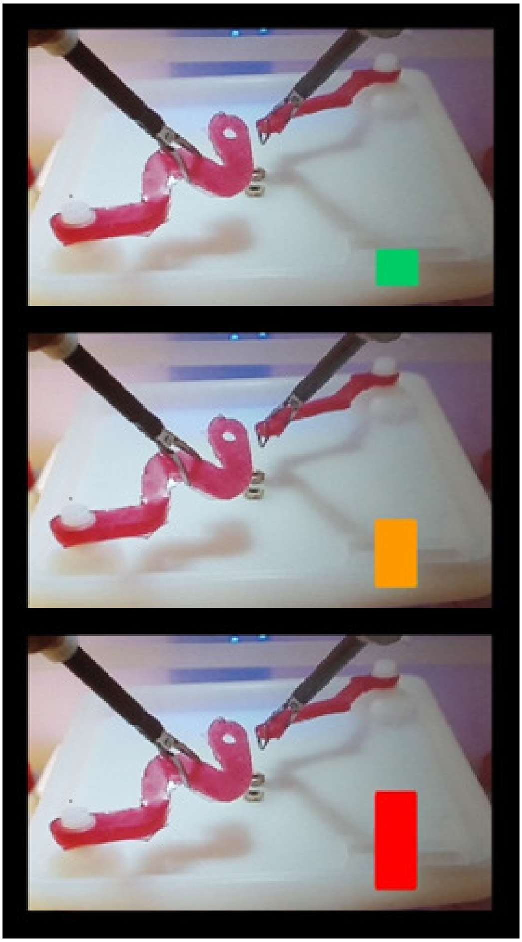


Figure 3. Display of Task 2 with force feedback during low (green), moderate (orange) and high (red) applied force.

Feedback design

To convey the force applied on the task, the visual force feedback design consisted of a vertical bar (Figure 3) that varied in size and colour as a result of the applied force on the task. A low amount of applied force was indicated by a small bar and similarly a high amount of force exerted on the task was indicated by a larger bar.

The colour of the force feedback bar was chosen consistent with existing preconceptions (Sigrist et al., 2013). The bar gradually changed colour bottom-up from green to yellow to orange to red depending scaled with the amount of exerted force. Warning triangles were presented in each corner of the display if extreme force was applied to prevent rupture of the strips. Since the elasticity of the artificial tissue (silicon) is close to that of uterus tissue the safety thresholds associated with uterus tissue were used in the colour scheme of the force feedback (Rodrigues et al., 2012).

Training protocol

First, participants signed an informed consent form and filled out a short demographics questionnaire. Next, participants familiarised themselves with the instruments, because understanding of equipment is important for safe laparoscopic surgery (van Hove et al., 2012). Prior to the pre-test, the participants were presented with visual on-screen instructions how to complete Task 1 (Figure 1). All participants were told to handle the tissues with care to prevent damage of the elastic components and to keep vision on the instruments at all time. After completing Task 1, instructions for Task 2 (Figure 1) were presented on the display. Participants performed Task 2 twice to create a reliable baseline. All participants performed Task 1 (placement of thread in flap) and Task 2 (connection of the silicon strips) during the pre-test without feedback of the tissue manipulation force.

Hereafter, all participants received instructions explaining the visual force feedback showed on the screen during training. As the type of force feedback during Task 2 was group dependent, this part of the explanation was different for each group. All participants were told that the training consisted of 6 trials of 5 minutes of Task 2. Participants were asked to complete Task 2 multiple times for the duration of each trial.

After the training, participants read the instructions for Task 1 again and were asked to perform the post-test (Task 1 and Task 2) without presentation of visual force feedback. A week later, all participants were asked to perform the follow-up test. The procedure was identical to the pre-test and post-test. After completing Task 1 once and Task 2 twice the participants received a certificate.

Performance parameters

Based on the proven classification power in earlier studies (Horeman, Dankelman et al., 2014), the parameters Maximum absolute force, Mean absolute nonzero force, Force volume and Max force area and (task completion) time were selected to establish a learning effect and to differentiate between the groups that trained with different types of feedback (Horeman, 2014).

Mean Absolute NonZero Force: The mean absolute force applied solely during application of force in Newton.

Maximum Absolute Force: The highest absolute force in Newton that was applied on the training task during the measurement.

Force Volume: If the force data is presented in 3D, three orthogonal principal components can be found indicting the three largest standard deviations of the force. The Force Volume is the volume of an ellipsoid fitted around those three standard deviations.

Max Force Area: If the absolute force is presented in time, the Max Force Area indicates the largest surface area under the graph. A force area is created between the moment in time the absolute force becomes larger than zero and the following moment in time the absolute force becomes zero again. Max Force Area units are presented in Newton second and referred to as peak force in earlier research.
Time: The time needed to complete the task, presented in seconds.

Statistics

Task 2 is used to identify differences between CFF, BFF and FFF on learning efficiency. To assure a valid pre-test, post-test and follow-up test data of Task 2, the mean of two measurements was taken. A paired sample t-test was used to compare the pre-test mean scores with the post-test mean scores of Task 1 and Task 2 separately. A paired sample t-test was also used to compare the pre-test mean scores and follow-up test mean scores of Task 1 and Task 2 separately.
Differences between the mean scores of the three groups in the pre-test, post-test and follow up post-test of Task 2 were examined using multiple one-way MANOVA's. Post hoc tests with Bonferroni correction were performed with a significance level of $p < 0.05$. Completion times data were non-normally distributed so we performed Wilcoxon Signed rank tests (for differences between pre-, post and follow up measurements) and Kruskal-Wallis tests for comparing between groups.

Results

Statistical differences between groups

To get insight into the effect of the feedback type on the force parameter results, the parameter outcomes of the training trials in relation to the pre, post and follow up trials are presented in Figure 4 (task 1) and 5 (task 2).

No significant differences on the force parameters were found between groups on task 1. The one-way MANOVA indicated no significant differences between the mean scores of the three groups in the pre-test on Task 2. Although the one-way MANOVA of Task 2 on the post-test revealed no significant multivariate main effect between groups, a significant univariate main effect was observed for the Mean Absolute NonZero Force (F (48, 2) = 4.303, p = .019, partial η2 = .152, power = .722) but not for the remaining force parameters. For this Absolute NonZero force, the Bonferroni post hoc tests showed a significantly lower mean score for the CFF group compared to the BFF group (p = .005). The one-way MANOVA performed on the mean scores of the three groups in the follow-up test did not reveal any significant differences between groups.

Kruskal Wallis tests showed no significant differences in completion times on task 1 and task 2 between the groups on the pre, post and follow up tests.

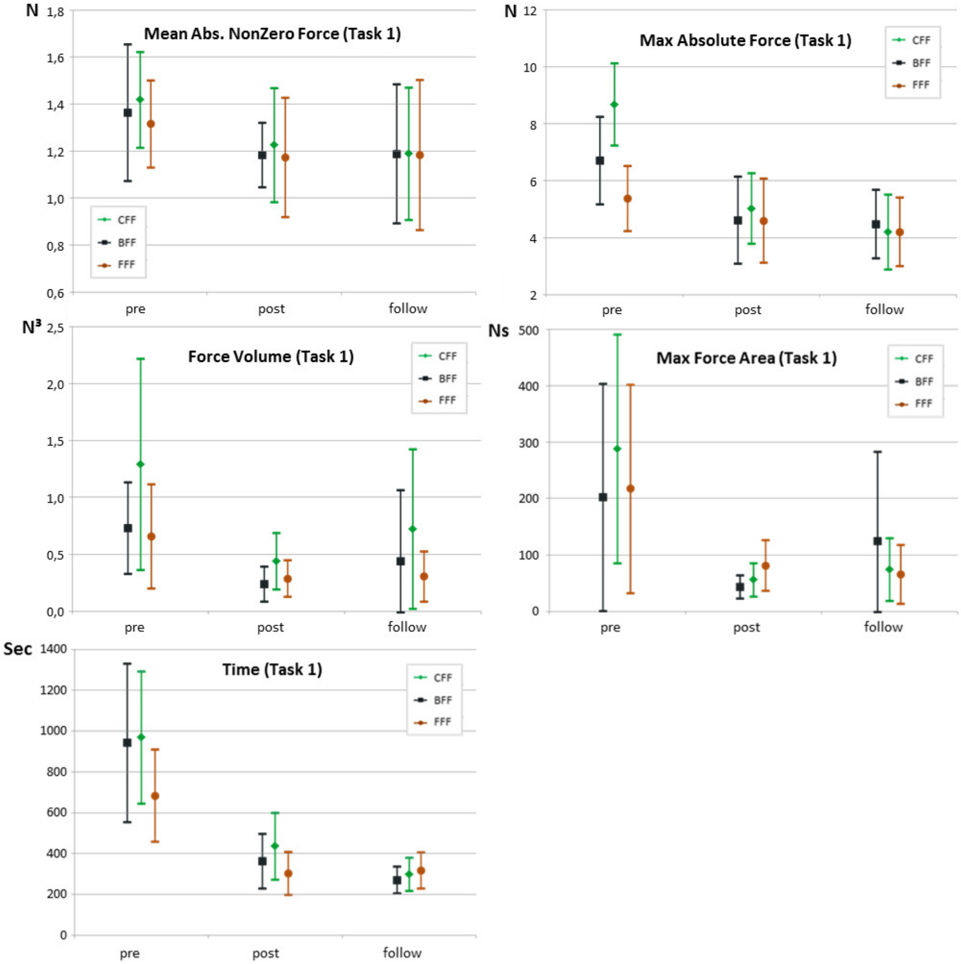


Figure 4. Mean scores for Mean Absolute Nonzero Force, Max Absolute Force, Force Volume, Max Force Volume and Time with 95% confidence intervals of the untrained task 1 divided for the three methods (CFF: Constant Force Feedback, BFF: Bandwidth Force Feedback, FFF: Fade-in Force Feedback).

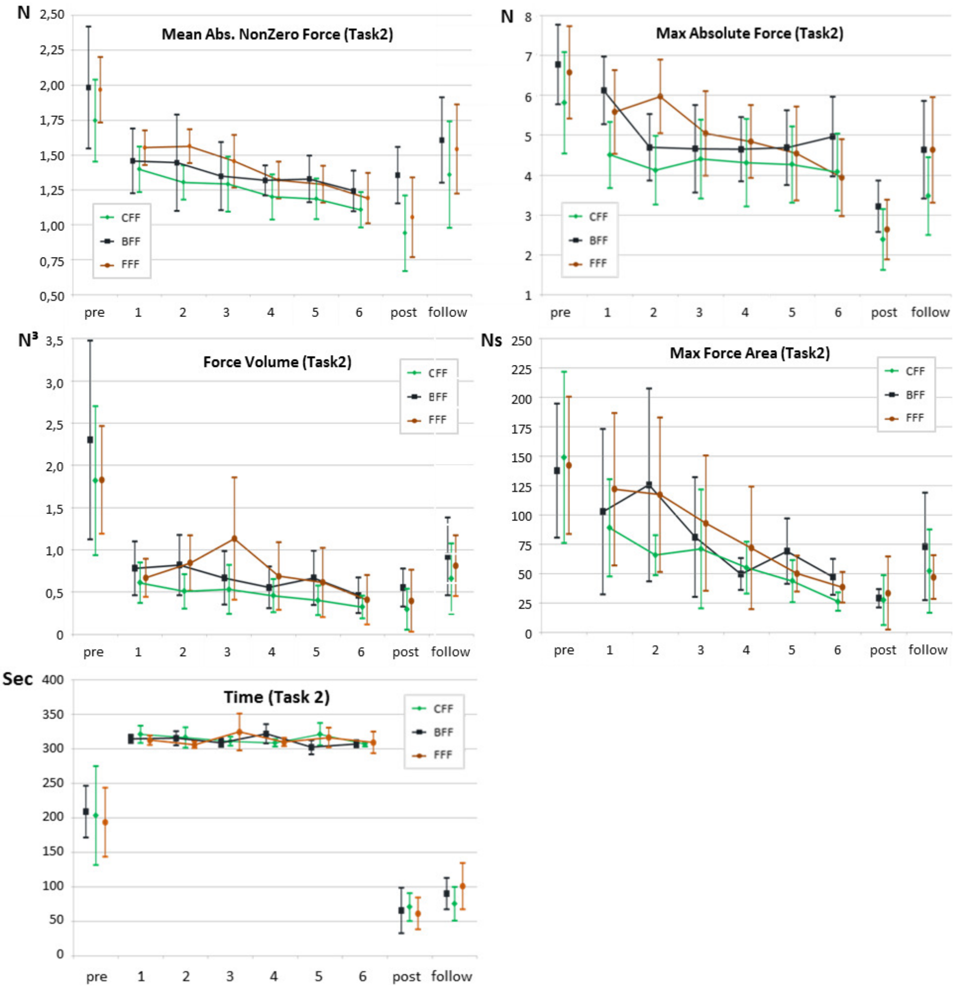


Figure 5. Mean scores for Mean Absolute Nonzero Force, Max Absolute Force, Force Volume, Max Force Volume and Time with 95% confidence intervals of the trained task 2, divided for the three methods (CFF: Constant Force Feedback, BFF: Bandwidth Force Feedback, FFF: Fade-in Force Feedback).

Differences between pre, post and follow up measurements

Task 1

Comparison of the pre-test mean scores with the post-test mean scores with a paired sample t-tests indicates that participants significantly decreased their applied Mean Absolute NonZero Force t(48) = 2.441, p = .018, Max Absolute Force t(48) = 5.866, p < .001, Force Volume t(48) = 3.446, p = .001 and Max Force Area t(48) = 3.419, p = .001. A week after the training, participants still applied significantly less Mean Absolute NonZero

Force t(48) = 2.02, p = .049 and Max Absolute Force t(48) = 5.809, p < .001, Force Volume t(48) = 2.479, p = .017 and Max Force Area t(48) = 2.692, p = .010.
Participants significantly reduced their completion times on task 1 from the pre-test to the post-test (Z = 5.217, p < .001) and from the pre-test to the follow-up test (Z = 5.784, p < .001).

Task 2

The paired sample t-tests indicated that the participants significantly decreased their applied force in the post-test compared to the pre-test. (Mean Absolute NonZero Force t(49) = 6.656, p < .001; Max Absolute Force t(49) = 11.057, p < .001; Force Volume t(49) = 6.187, p < .001; Max Force Area: t(49) = 6.153, p < .001). Furthermore, when comparing the pre-test mean scores to the follow-up test mean scores of Task 2, we find that the participants were able to significantly decrease their applied Mean Absolute NonZero Force, t(48) = .004, p = .004; Max Absolute Force, t(48) = 5.321, p < .001; Force Volume, t(48) = 4.633, p < .001 and Max Force Area, t(48) = 4.427, p < .001.
For completion times on task 2, participants reduced their scores from the pre-test to the post-test (Z = 5.720, p < .001) and from the pre-test to the follow-up test (Z = 5.436, p < .001). There was a minor increase in median completion times from the post-test to the follow-up test (Z = 3.065, p < .001).

Discussion

The aim of this study was to determine whether different visual force feedback types (i.e., constant, bandwidth and fade-in) have different effects on the learning curve when acquiring tissue manipulation skills. Only the force parameter Mean Absolute NonZero Force showed significantly lower mean scores for the BFF group compared with the other groups. This lack of meaningful differences between the groups in the follow-up test seems remarkable because of the difference in total time that participants received visual force feedback in the three groups.
Comparing the learning curve trajectories of the three groups provides insight in the impact of visual force feedback on the force parameters that reflect dangerous tissue handling (i.e. Max and mean NZ force and Force area). Participants in the FFF group applied relative high force in the first two trials, in comparison to the other groups. When the force feedback became more prevalent in the remaining trials, participants in the FFF group managed to improve their tissue manipulation skills in a faster rate until the level of the participants in the other groups was reached. This shows potential for more advanced tasks as it allows the trainee to decide to master basic skills (instrument handling, fulcrum effect, bimanual cooperation, etc.) first before focusing on tissue handling aspects.

Although all feedback types seem to work effectively for the performed tasks, the Bandwidth Force Feedback is the only type of feedback that respects individual progress (Sigrist et al., 2013). It therefore minimizes the duration of visual force feedback presentation while similar performance improvements are observed. This indicates that brief exposure to visual force feedback at the right moment in training is already sufficient to decrease the applied force.

Observing the results in general, one can clearly identify learning curves for all of the force parameters on the trained task. All participants significantly decreased their mean scores on all force parameters compared to the pre-test. After one week, a clear training effect was still prevalent since participants performed the trained task with significantly lower mean scores on all of the force parameters compared to the pre-test. Prospects of the training method are promising because laparoscopic tissue manipulation skills acquired in one and a half hour are still retained after a week. In addition, transfer to a different task with dissimilar characteristics is observed as well. Participants were able to significantly decrease their scores on all force parameters on a dissimilar untrained task. In the follow-up test, participants had significantly lower mean scores on the untrained task on aforementioned force parameters. The experimental training groups aside, one can conclude that the training method with visual force feedback is generally effective in decreasing the applied force.

Limitations

Horeman, van Delft et al. (2014) have previously shown that participants significantly decreased their applied force when presented with constant visual force feedback compared to a control group were no feedback was given. This study aimed to tune the visual force feedback training method, therefore the control group in this study was a group with constant feedback. The lack of a no visual force feedback group can be seen as a limitation of the current design.

Another limitation is the extended period of training on one task. Multiple participants reported to be bored as a result of the lengthy training trials. Usually, such emotional states can cause demotivation and decrease task engagement (Fisherl, 1993). Ultimately this could have resulted in a decreased potential to acquire the laparoscopic tissue manipulation skills.

Not using a power calculation to determine the required group size can be seen as a limitation. Instead, the study of Horeman, van Delft et al. (2014) was used to determine the absolute minimum group size required to distinguish the most important differences in performance. The maximum actual size was determined by the number of participants willing to collaborate.

Recommendations

The study shows that training effects of the ForMoST device in combination with the presentation of visual force feedback are retained for at least a week. Second, these training effects also transfer to an untrained task with other characteristics. It is of utmost importance that the acquired laparoscopic skills can be transferred to the real occupational setting as well. Proving predictive validity would increase the legitimacy of this training method (Mané, Adams & Donchin, 1989). Further research is required to understand if, and to what extent, the acquired laparoscopic skills are transferable to the OR. Reassessment on hybrid box trainers at a later point in time should also clarify the long-term retention of the acquired laparoscopic skills. Participants should be reassessed after an extended interval to reveal the effectiveness of the training method over time (Hiemstra, 2012).

Of main importance for the student surgeons is to acquire laparoscopic tissue manipulation skills, which includes awareness of the consequences of too much applied tissue force and the level of their tissue interaction force. The training method that is used in this study supports the participant in acquiring those skills and should therefore be included in the laparoscopic surgical training curriculum. Adding requirements for force parameters scores in the performance assessment of residents will ensure surgeons possess better laparoscopic skills after completing training.

Conclusion

All visual force feedback groups showed to be equally effective in decreasing participants applied task force. The learning curves recorded in training, the mean scores of the force parameters in post-test and the retention effects after a week indicate that training with visual force feedback results in enhanced laparoscopic tissue manipulation skills. As the Bandwidth Force Feedback type is only present when force levels are dangerous, it minimizes attentional distraction and is therefore preferable for training.

Acknowledgements

The authors would like to thank MediShield BV for providing laparoscopic instruments, training tasks and laparoscopic box trainer. The authors like to thank Dennis van Gerwen from the TU Delft for helping the authors with programming of the visual force feedback.

CHAPTER 06

VARIED PRACTICE IN LAPAROSCOPY TRAINING: BENEFICIAL LEARNING STIMULATION OR COGNITIVE OVERLOAD?

Edward N. Spruit MSc.
Luca Kleijweg MSc.
Dr. Guido P.H. Band
Prof. Jaap F. Hamming

Based on the publication:
Spruit, E. N., Kleijweg, L. M., Band, G. P., & Hamming, J. F. (2016).
Varied Practice in Laparoscopy Training: Beneficial Learning Stimulation or
Cognitive Overload?. *Frontiers in Psychology*, 7, 685.

Abstract

Determining the optimal design for surgical skills training is an ongoing research endeavor. In education literature, varied practice is listed as a positive intervention to improve acquisition of knowledge and motor skills. In the current study we tested the effectiveness of a varied practice intervention during laparoscopy training. 24 trainees (control group) without prior experience received a three week laparoscopic skills training utilizing four basic and one advanced training task. 28 trainees (experimental group) received the same training with a random training task schedule, more frequent task switching and inverted viewing conditions on the four basic training tasks, but not the advanced task. Results showed inferior performance of the experimental group on the four basic laparoscopy tasks during training, at the end of training and at a two month retention session. We assume the inverted viewing conditions have led to the deterioration of learning in the experimental group since no significant differences were found between groups on the advanced laparoscopic task (the only task that had not been practiced under inverted viewing conditions). Potential moderating effects of inter-task similarity, task complexity and trainee characteristics are discussed.

Introduction

Acquisition and retention of complex perceptual and motor skills is often perceived as a challenge in various fields (aviation, driving, sports, medicine). With constant improvements in modern technology, more specialists require thorough and adequate training in order to operate new devices effectively. With increased popularity of minimally invasive procedures in health care, various medical professions require a new level of technical skills. For appendectomy and cholecystectomy, laparoscopy has become the default technique of operating. Minimally invasive procedures provide potential and proven benefits for patients (smaller scars, shortened hospitalization, faster return to daily activities), but are more difficult to learn for trainees. In laparoscopic surgery, practitioners need to attain perceptual and motor skills required to deal with the increase in complexity of the task at hand. Diminished tactile feedback, loss of depth perception cues, fewer degrees of motion freedom, amplified tremor and counter-intuitive instrument movement (Gallagher & O'Sullivan, 2012) raise the inherent difficulty of the task.

Given these challenges, there has been an increase in demand for quality training of surgical residents before they start practicing in the operating room. In order to ensure patient safety, it is important that residents are prepared thoroughly and have a high retention and transfer of laparoscopy skills by the time they perform their first surgical procedure. There is a clear interest in both education and research to investigate best practices to conduct medical training (Gallagher & O'Sullivan, 2012; Spruit, Band, Hamming & Ridderinkhof, 2013).

Variability of practice and frequent switching of different training tasks is a common suggestion in literature for efficient training (Schmidt & Bjork, 1992). However, it is thus far unclear whether variability of practice is beneficial in every context of training and to what extent the effect can be generalized (e.g., Brady, 2004; Barreiros, Figueiredo & Godinho, 2007; Hall, Domingues & Cavazos, 1994; Jones & French, 2007). An important research question facing trainers is: How much variability is ideal during training of complex motor skills?

The Contextual Interference Effect

The contextual interference (CI) effect (Battig, 1966; Shea & Morgan, 1979) is characterized by an impairment in performance during the skill acquisition phase of training, brought about by an increase of variability of training task parameters. When the training context changes frequently, the trainee experiences more interference of his motor actions, since different tasks require different actions. This finding is prevalent in the domain of memory (Schneider, Healy & Bourne, 2002), but has also been widely tested in the field of motor skill acquisition (Magill & Hall, 1990). Examples of such variability in the motor domain are randomization (as opposed to block wise presentation) of the order of training tasks or frequently adjusting the conditions of a task (i.e. adjusting the distance to the target, the size of the ball or target in a basketball free-throw training task). Increasing the amount of CI tends to increase long-term retention of performance and flexibility of trained skills. Both are valuable for surgical residents since the interval between training and application on the job can be substantial and many different cases of anatomy can occur during surgical operations.

Two hypotheses are considered as an explanation for the effect (Boutin & Blandin, 2010). The elaboration hypothesis states that randomization of training tasks facilitate inter-task comparison during processing by the trainee, which leads to a more elaborate representation of different aspects of each training task. The reconstruction hypothesis suggests that due to the nature of more frequent task switching on a randomized training schedule, trainees are forced to rebuild the representation of their motor strategy whenever a different task is presented to them. In other words, they continually have to update their modus operandi whenever the training task changes, which strengthens the motor representation in long-term memory. It also helps trainees to stay engaged in learning.

The CI effect is similar to the spacing effect (Donovan & Radosevich, 1999), in that both center on the principle of increasing the effectiveness of learning episodes by splitting up training time into separate and different training intervals. When spacing training, this is achieved by utilizing multiple training sessions, whereas varied practice does so by creating more distinct training segments within one session. Both interventions force the trainee to update their working memory with instructions and task strategies more frequently than on a blocked or massed training schedule. An important difference is that spacing reduces the amount of retrograde interference (the extent to which a second training task practiced directly after a first training task compromises learning and retention of the first task, Brashers-Krug, Shadmehr, & Bizzi, 1996; Shea & Graf, 1994), whereas training with increased variability induces interference.

Both spacing and varied practice capitalize on the benefits of maximizing working memory recruitment throughout the learning process, which helps to keep trainees engaged in training. It is unclear whether this extra interference is beneficial or it becomes a heavy load on an inherently complex task, such as laparoscopy. Cognitive load theory (Paas, Renkl & Sweller, 2003) rests on the assumption that learning is optimal when trainees have enough spare cognitive resources available to reflect on what they are learning (germane load), next to the resources required to perform the task (intrinsic load) and the resources needed to comprehend instructions and filter noise (exogenous load). Hence, adding variable conditions to a complex task such as laparoscopy may lead to excess cognitive load (van Merriënboer, Kester & Paas, 2006) and result in a skill—challenge imbalance and a disruption of flow in the learning experience (Engeser & Rheinberg, 2008). Intrinsic germane load can be increased by adding more CI (more variability) to a practice schedule (van Merriënboer & Sweller, 2010).

Brady (2004) found that the CI effect is more pronounced in basic experimental research, yielding moderate effect sizes, but small effect sizes are found for applied studies. A possible explanation is that most applied settings (like sports) already contain high variability even in blocked training conditions, whereas experimental lab settings with basic motor tasks are better able to control for variability of conditions. Also, basic laboratory research allows for rigorous control of variables, which have a more natural flow in applied settings (Barreiros, Figueiredo & Godinho, 2007). The current setting (laparoscopy simulator training) leaves us somewhere in the middle between a lab and applied setting.

The amount of research on varied practice and contextual interference in medical training is limited, although it is a frequently suggested learning principle based on research in other fields (Schaverien, 2010; Spruit, Band, Hamming & Ridderinkhof, 2013).

In the field of surgical training, Brydges, Carnahan, Backstein and Dubrowski (2007) provided training on a bone-plating surgical task on three different practice schedules (whole-task training, blocked part-task training and random part-task training) and found no support for the CI effect, since the whole-task training group showed more improvement than the random part-task training group. However, these results ought to be interpreted with caution since individual difference scores were used in order to account for group differences in baseline scores, rather than comparing post-test scores. Other research on laparoscopy training has shown that alternating the visual presentation during laparoscopy training increases the rate of automatization to the fulcrum effect (Jordan, Gallagher, McGuigan & McClure, 2000).

The aim of the current study is to investigate the effect of varied practice in laparoscopy training. Based on the CI literature, we first hypothesize that performance of the experimental (varied practice) condition will be impaired during and immediately following the acquisition phase of training, whereas performance will be enhanced at a retention session planned two months after training. Based on cognitive load theory, our second and competing hypothesis states that performance will be impaired in the experimental condition and the control condition will have superior performance on all sessions. We also expect all participants to improve in performance on all tasks from baseline to the end of training and retention, regardless of the condition.

Method

Participants

60 medical and psychology students without prior experience in laparoscopy training were enrolled in the study (data in prior studies yielded no significant differences in laparoscopic performance and aptitude between medical and other university students, hence the inclusion of psychology students in the sample). Fifty-two students (39 female, 43 medical), all right-handed, age ranged from 18-29y (mean = 21.54) completed all sessions. After completing the training, all participants received a certificate as a reward for participating in the study.

Apparatus

Training centered primarily on practicing skills on a laparoscopic box trainer. Students practiced four basic and one advanced task with previously established construct validity (Kolkman et al., 2008). These tasks train perceptual and motor skills (instrument handling, depth perception, adjusting to the fulcrum effect) that are key to proficiency in laparoscopic surgery. The first task requires participants to stretch a rubber band around a set of twelve spikes. In this task a trainee learns to work with forces. In the second task, participants string a pipe cleaner through a set of four rings. This task aims to train bi-manual dexterity. The third task involves the placements of small beads on a pegboard and requires high precision of motor actions. In the fourth task, a circle should be cut in a rubber glove, which trains participants in exposure and dissection skills. In the advanced task, participants train the skill of intra-corporeal suturing. In the current study, participants were taught how to create three knots, starting with the needle in their right instrument, using two throws for the first knot. One throw was used for the second knot starting in the left instrument and one throw for the third knot starting from the right instrument (see the reference list for an online appendix with videos of all our laparoscopic training tasks[*]). To prepare participants for the advanced task, an open model of suturing was used in order to familiarize them with the procedure of knot tying before attempting the task in the box-trainer.

Performance of participants on the box-trainers was recorded on a connected PC with the use of a video splitter and grabster (Terratec Grabster AV 400/AV300 MX) to convert video output to separate video files via USB. The USB signal was converted to .mpg files by VLC Media Player for Windows.

Self-report questionnaires were used to acquire data on sex, age, dominant hand, academic year, prior sports, music, and gaming experience, goal orientation (Seijts & Latham, 2006) and growth mindset (Dweck, 2006). This data was collected to test comparability in these variables between the two conditions.

Training Programs

Participants were quasi-randomly assigned to the experimental (varied practice) condition (n=28) or the control group (n=24). Depending on their availability, participants were scheduled for three training sessions on a specific week day for three consecutive weeks. Both groups participated in three training sessions of two hours per session. Each session consisted of sixty minutes of practice, twenty minutes of instructions, ten minutes of break and a thirty minute period during which each participant's performance was recorded. During the first week (training session I), participants practiced the four basic laparoscopic tasks and suturing on the open model in order to learn the basics of suturing and knot tying. In the second week (training session II), intra-corporeal suturing was introduced and participants practiced all five laparoscopic tasks during the training sessions (II and III) in week two and three. Participants returned to the lab after two months for a retention session, during which performance was recorded without any prior practice on that session. Participants received standardized instructions from the trainers (first and second author) and a self-directed feedback (Strandbygaard et al., 2013) protocol was used: the majority of instructions were provided by video and feedback to trainees was minimized, and only provided if initiated by the trainees.

Experimental and Control Condition

In the control condition, participants practiced each task in a set order, starting with practice on the easier tasks and progressing to the more difficult tasks. Each task was practiced for one time interval per session (see Table 1). The experimental group practiced each task for multiple smaller time intervals per session and switched more frequently between training tasks during each session than the control group. However, the total training time per task was equal for both conditions. Five box trainers with a training task were prepared and the trainees switched every six (training session one) or three minutes (training session two and three) between them when the experimenter indicated. The schedule for the varied practice group was determined randomly.

Also, participants in the experimental condition practiced each of the four basic laparoscopic tasks under inverted viewing conditions for half of the training time of these tasks (see Table 2). Inverted viewing conditions were achieved by flipping the laparoscopic box trainer by 180 degrees across the front-back dimension and creating extra insertion points for the instruments on the other side of the box-trainer. This setup implied that the camera was now facing towards the trainee, which inverted the trainee's perception of the movements made with the laparoscopic instruments.

Table 1. Practice schedule of training session I for both conditions.

Blocked Condition		Contextual Interference Condition	
Activity	**Minutes**	**Activity**	**Minutes**
Welcome, planning & informed consent	5	Welcome, planning & informed consent	5
Instruction videos tasks 1-4	10	Instruction videos tasks 1-4	10
Recording of baseline tasks 1-4 and instruction on knot tying	30	Recording of baseline tasks 1-4 and instruction on knot tying	30
Practice task 1 (rubber band)	12	Practice round 1 (5x6 minutes on each of the 5 practice tasks)	30
Practice task 2 (pipe cleaner)	12		
Practice task 3 (placing beads)	12		
Break	10	Break	10
Practice task 4/5 (cutting circle/ knot tying)	12	Practice round 2 (5x6 minutes on each of the 5 practice tasks)	30
Practice task 4/5 (cutting circle/ knot tying)	12		
Debriefing	5	Debriefing	5

Table 2. Task order of the contextual interference group training session 1

Round 1	Round 2
Placing Beads	Pipe cleaner
Inverted rubber band	Inverted cutting circle
Inverted pipe cleaner	Rubber band
Cutting circle	Inverted placing beads
Knot tying	Knot tying

Under normal viewing conditions the camera was facing away, in line with the trainee's field of vision. Each training task, trainees were unaware of which setup (inverted or regular viewing conditions) they were faced with until they started to practice. We chose this way of inverting perception (changing camera orientation) over other means (mirroring the screen monitor) because this setup is more congruent with a real surgical setting, where the angle of the laparoscope is frequently different from the angle of the surgeon's field of view.

Outcome Measures

The first and second author scored task completion times based on video recordings. Recordings at the beginning of the first session served to assess the baseline skill level. Performance was also measured at the end of the second and third training sessions. In order to test the durability of acquired skills, we assessed the participants' retention two months after training, without any practice beforehand. In total, there were four moments of measurement.

Statistical Analysis

Data distribution histograms were analyzed and Shapiro-Wilk tests were performed on the data to check for normality. We tested whether groups were comparable at baseline in terms of age, sex, dominant hand, academic year, academic specialization, musical, gaming, sports activity and personality factors.

For each of the five training tasks, Mann-Whitney tests were done to test whether differences were present between groups at each stage of training (baseline, training session two, training session three, and retention). Wilcoxon signed-rank tests were done to check if improvements within trainees occurred between subsequent training sessions. We performed additional analyses (non-parametric correlations) to explore whether there were any interesting relations between questionnaire data and performance on the laparoscopic tasks.

Statistical significance was determined at $p < 0.05$, one-tailed for within-condition progress, two-tailed for between condition comparisons.

Results

Fifty-two students ($N_{exp} = 28$, $N_{con} = 24$) completed all sessions and were included in data analysis. As predicted, data were non-normally distributed, so non-parametric tests were used.

If trainees were unable to complete a task within ten minutes, a score of 601 seconds (a score that would automatically be assigned as the highest rank in the non-parametric tests) was assigned. This was done to avoid selective drop-out from our sample based on poor performance in the cases where trainees were unable to complete a task. This was the case for 31 out of 988 moments of measurement.

Baseline Check

Mann-Whitney tests revealed there were no differences between conditions in gaming and sports activity, although the experimental group (Mdn = 0 out of 5, IQR = 0) practiced slightly less with musical instruments than the control group (Mdn = 0 out of 5, IQR = 2), $U = 247.5$, $z = -2.082$, $p = .041$, $r_{rb} = -0.29$. The two conditions were similar in age ($Mdn_{exp} = 21y$; $Mdn_{con} = 22y$), $p > .05$. Independent samples t-tests showed comparable distributions on goal orientation and growth mindset among conditions.

Chi-square tests showed that the control group contained significantly more male participants (42%) than the experimental group (11%), $p = .022$. Also, the experimental group contained more psychology students (29%) compared to the control group (4%), $p = .028$. Male participants also performed better on most of the laparoscopic tasks (17 out of 19 measurements), but this difference was only significant in one instance (beads task retention, p = .01). We controlled for these factors after our main analysis.

On the first three basic laparoscopic tasks (elastic band, pipe cleaner, beads) there were no significant differences in baseline scores between conditions (see Table 3). For the circle cutting task, a significant trend in favor of the control group (p = .001) was found at baseline.

Within-group progress across training sessions

The median scores at baseline, the end of training Sessions II, the end of training session III and at retention for the first four laparoscopic tasks are shown in Table 3. Participants in both groups significantly improved on all training tasks during training and showed deterioration from the end of training Session III to the retention session on the pipe cleaner task (experimental condition), cutting circle task (control condition) and the intra-corporeal suturing task (both conditions).

Main Analysis: Between-condition comparison

In Table 3, median scores of both conditions are compared for each task at each training session. Participants in the control condition performed significantly better at the basic tasks at Session II, Session III and retention, with the exception of beads (training session III) and circle cutting (retention). For the intra-corporeal suturing task, no significant differences between the conditions were found. Estimates of effect sizes show moderate effects for the first four basic laparoscopic tasks at Session II, Session III and retention, but no effects for intra-corporeal suturing.

By standardizing and compounding the scores of the four basic laparoscopic tasks, this difference is clearly illustrated (see Figure 1). For the advanced laparoscopic task (suturing), the conditions have comparable scores (see Figure 2), although a trend can be observed where the median for the experimental condition is higher than the median of the control condition at the end of training Session II, roughly similar at the end of training Session III, but lower at retention.

Extended analysis and confound check

Given the lower scores for the control group at baseline on the circle task and a similar trend for the other basic tasks, we wanted to test whether these may have confounded our analysis. To achieve this, a post-hoc case-controlled analysis was done for the standardized total baseline scores. We matched the two conditions in standardized total baseline scores by excluding the lowest scoring participants from the control condition and the highest scoring participants from the experimental condition until the two conditions had roughly equal mean rank scores in a Mann-Whitney test. The final sample for this analysis was 46 (N_{exp} = 24, N_{con} = 22). We found no significant differences relative to the results of our main analysis, although the observed trend for intra-corporeal suturing changed in favor of the experimental condition on all sessions (Training Session II: Mdn_{exp} = 386.5 à 303.0; Mdn_{con} = 359.0 à 393.5) (Training Session III: Mdn_{exp} = 192.5 à 179.5; Mdn_{con} = 189.5 à 199.5) (Retention Session: Mdn_{exp} = 243.0 à 220.5; Mdn_{con} = 261.5 à261.5), but did not reach significance (p = .216; p = .386; p = .425; respectively). Matching groups for baseline scores on just the cutting circle task, musical activity, sex or study did not substantially alter our results.

Table 3. Results of the Mann-Whitney tests with median completion times (in seconds) and effect sizes (r_{rb}) for the five training tasks (horizontally).

	Mdn_{exp}	Mdn_{con}	U	Z	p	r_{rb}
Rubber band task						
Baseline	146.0	133.0	320.0	-.29	.78	-.04
Training Session II	63.0 $\downarrow^{***}$	41.0 $\downarrow^{***}$	136.0	-3.67	<.01	-.51
Training Session III	57.5 $\downarrow^{*}$	35.0 $^{\approx}$	202.0	-2.46	.01	-.34
Retention	58.0 $^{\approx}$	44.5 $^{\approx}$	209.0	-2.33	.02	-.32
Pipe cleaner task						
Baseline	141.5	123.5	288.0	-.88	.38	-.12
Training Session II	63.0 $\downarrow^{***}$	42.0 $\downarrow^{***}$	143.5	-3.54	<.01	-.49
Training Session III	54.0 $^{\approx}$	43.5 $^{\approx}$	198.5	-2.53	.01	-.35
Retention	66.0 $\uparrow^{*}$	42.5 $^{\approx}$	108.0	-4.19	<.01	-.58
Beads task						
Baseline	310.5	261.5	261.0	-1.38	.13	-.19
Training Session II	172.5 $\downarrow^{***}$	121.5 $\downarrow^{***}$	167.5	-3.09	<.01	-.43
Training Session III	147.5 $^{\approx}$	119.5 $^{\approx}$	240.0	-1.76	.08	-.24
Retention	158.5 $^{\approx}$	126.5 $^{\approx}$	204.0	-2.42	.02	-.34
Cutting circle task						
Baseline	428.0	303.5	166.0	-3.13	<.01	-.43
Training Session II	183.5 $\downarrow^{***}$	130.0 $\downarrow^{***}$	224.5	-2.05	.04	-.28
Training Session III	139.0 $^{\approx}$	107.5 $\downarrow^{**}$	221.5	-2.10	.04	-.29
Retention	116.5 $^{\approx}$	118.5 $\uparrow^{*}$	333.5	-.05	.97	-.01
Intra-corporeal suturing task						
Training Session II	386.5	359.0	333	-.06	.96	-.01
Training Session III	192.5 $\downarrow^{***}$	189.5 $\downarrow^{***}$	322	-.26	.80	-.04
Retention	243.0 $\uparrow^{**}$	261.5 $\uparrow^{*}$	322.50	-.25	.81	-.03

Within-subject progress is indicated vertically for training session II, III and the retention session compared to the previous session ($^{\approx}$ for no significant difference; $\downarrow$/ $\uparrow$ for significant improvement/deterioration, * = p < .05; ** = p < .01; *** = p < .001)

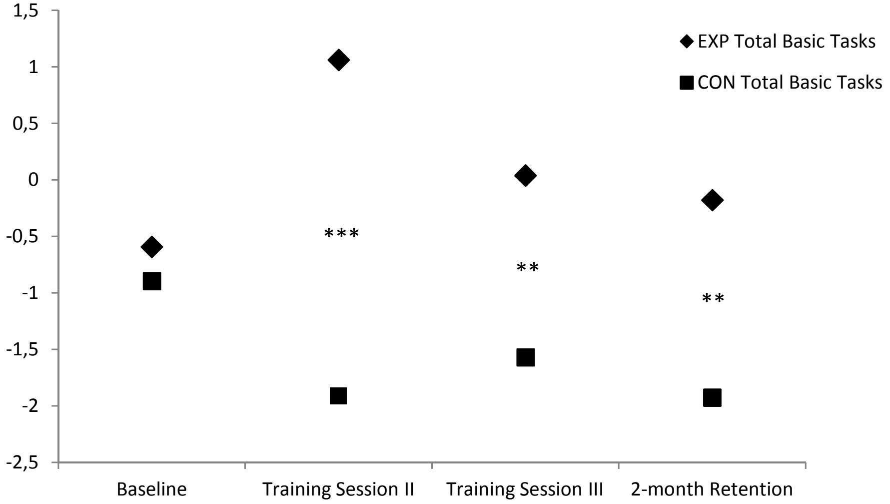


Figure 1. Standardized total scores of the first four basic laparoscopic tasks at end of the second and third training session, and at retention for both training groups (Between-subject differences: * = p < .05; ** =p < .01; *** = p < .001).

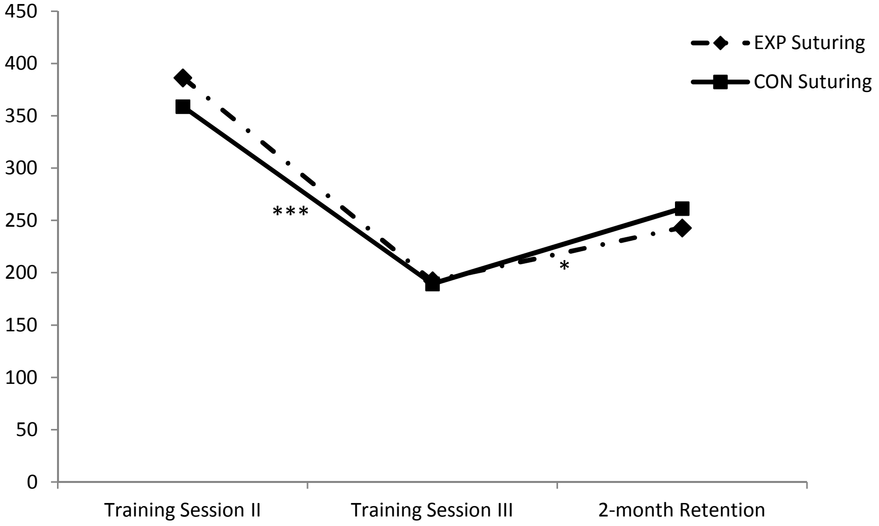


Figure 2. Median completion times (in seconds) for the advanced laparoscopic task at the end of the second and third training session, and at retention for both training groups (Within-subject progress: * = p < .05; ** =p < .01; *** = p < .001).

Discussion

The results in the current study show a clear impairment of performance in the experimental group with varied practice on four basic laparoscopic tasks during training and at a retention session two months after training. Although the tempered performance on the short-term (session two and three) during training is characteristic of the contextual interference effect (Brady, 2004), the experimental condition did not yield superior long-term performance (retention). Since retention performance is a better indicator for learning (Schmidt & Bjork, 1992), the additional contextual interference in the experimental condition has likely impaired the learning process of the participants, which supports the second hypothesis (based on cognitive load theory).

What is notable, however, is that the scores on intra-corporeal suturing were not significantly different in the control and experimental group. Since this was the only task that was not practiced in inverted viewing conditions, it leads one to conclude that the inverted conditions have had a more detrimental effect as compared to the random training task schedule and the more frequent task switching. It is also possible that the random schedule had a positive effect, but it had been canceled out by the impaired learning caused by inverted viewing conditions. It is unclear whether or not the randomized schedule and frequent task switching has had a positive effect or not, but from the results of the intra-corporeal suturing task we can derive that it has not been a detrimental intervention.

Inter-task similarity and contextual interference

There are different ways of introducing contextual interference and this allows for various ways in designing training. Switching between tasks that use a broader range of motor functions is known to lead to different types of contextual interference (Magill & Hall, 1990). Switching between the pipe cleaner task and a rubber band task is not that big of a difference, since the tasks share a lot of similarities (same instruments, same perception, similar motor patterns). Different variations of the same type of motor task invoke the contextual interference effect less frequently than switching between motor tasks that use different motor programs.

For instance, we could have had participants switch between a laparoscopic training task and a venipuncture procedure training task or a completely unrelated motor task (like a basketball throw) altogether. Switching between tasks that use different motor programs creates a more difficult learning setting, which facilitates contextual interference (Magill & Hall, 1990). It is questionable whether such additional interference would have improved learning in the experimental condition in our current study given the high level of interference produced by the inverted viewing conditions. Nonetheless, it would be an interesting avenue for future studies and inter-task similarity is an important variable to consider when designing training of complex motor skills.

Task complexity

The tasks used in the current study are bound to a set of perceptual limitations. The notion that depth perception and tactile feedback are compromised makes for an inherently complex training task. Participants have to get accustomed to the fulcrum effect (Gallagher, McClure, McGuigan, Ritchie & Sheehy, 2008), the fact that whenever they move their hand in one direction, the tip of the corresponding instrument moves in the opposite direction. Task complexity was further increased for the experimental condition by adding the inverted viewing condition, where the fulcrum effect was cancelled out. Switching between these (fulcrum effect/no fulcrum effect) conditions induced substantial contextual interference during training.

In an earlier study, alternating between such conditions led to faster automatization to the fulcrum effect (Jordan, Gallagher, McGuigan & McClure, 2000). The criterion task in the mentioned study was not very complex (making incisions in an A4 paper) and only the horizontal axis was inverted by mirroring the monitor screen, rather than using the laparoscope from a different viewing angle. This results in just an inversion of left and right, as opposed to left and right and front and back, as was the case in the current study. From our results, we conclude that the amount of contextual interference may have led to a cognitive load that was too high for most of the participants. We suggest that the amount of contextual interference has to be carefully gauged based on a trainee's current proficiency level in order to attain the desired learning outcomes. The gauging principle also applies to the training context as stress can also have an impact on the amount of cognitive load a trainee can handle. This is especially relevant in fields such as aviation and medicine when simulating emergency scenario's (Taber, 2014).

Trainee characteristics

Novice trainees are impacted differently by contextual interference than more experienced trainees (Brady, 2004). In the current study participants had no prior experience with laparoscopy training. During the learning process, the same training task becomes less challenging over time as the trainee practices and gains more skill. During training, the cognitive load it takes to execute a complex task remains stable over time (intrinsic load) and the surrounding conditions (quality of instructions, noise, task schedule) under which the task is performed remains relatively stable over time as well (exogenous load), but the cognitive load of learnable aspects of the task (germane load) diminishes as skill accumulates (Sweller, van Merriënboer & Paas, 1998). A novice without any prior experience on the task will be more taxed by germane load than the same novice at the end of a three week training. A plausible explanation for the current findings is that participants in the experimental condition received an increase in task complexity (inverse viewing conditions) and exogenous load (task switching) that was too high given the capacity of the cognitive load they were able to handle. Learning the laparoscopic tasks provided substantial germane load and the result of the varied practice intervention led to a training that became challenging to the extent of impairing, rather than facilitating learning.

In future designs, an option would be to test a gradual increase in contextual interference, as is suggested by Porter and Magill (2010). This principle could be coupled with adaptive training, since no trainee's skill level is the same when they initiate training and when they finish training. The amount of challenge is subjective to the trainee and task complexity and contextual interference should be coupled with the skill of the trainee (Guadagnoli & Lee, 2004). Instead of burdening novices with too much contextual interference, trainers can opt to introduce inverted viewing conditions to intermediate and advanced trainees that can handle the additional cognitive load. Testing this intervention at a later stage of training may allow trainees to accustom to the different angles of the laparoscope that they will likely face in the operating room.

Study limitations, generalizations and conclusions

In the current study, participants were only measured on training tasks under regular viewing conditions. This may be seen as a limitation, since the control condition had more time-on-task practice during these circumstances as compared to the experimental condition. Measuring performance under inverted viewing conditions would alleviate this issue, but it is likely that this would be a serious challenge for the control condition, where this is a completely new context. Also, it is still unclear whether or not the trend observed for the suturing task was a positive contextual interference effect of the frequent task switching or that it was due to sampling error. Future studies could investigate a two-by-two design with the inclusion and absence of the two training interventions (frequent task switching and alternating viewing conditions) to further clarify whether frequent task switching is truly beneficial for learning laparoscopic skills.

At first glance the current results may seem indicative of favoring the training design of the control condition, given their superior performance at the end of training and at retention. We emphasize, however, that task complexity needs to be taken into account when designing training. Also, low variability training conditions are typically associated with lower levels of transfer (van Merriënboer et al., 2006) and examples exist in the literature where higher CI conditions during simulation training lead to higher transfer when training less complex motor tasks (like tennis; Broadbent, Causer, Ford & Williams, 2015). Performance on a training simulator does not necessarily equal performance in the operating room (transfer). The laparoscope can be positioned in different angles throughout a laparoscopic procedure in the operating room and the contents of the human abdomen are more complex than a basic laparoscopic box-trainer. We encourage trainers to try to incorporate this principle in their training and research, so that less learning has to take place in the operating room and more learning can occur during training in the lab. Based on our findings, we would urge trainers to be cautious initially when introducing new varied practice interventions to inexperienced trainees, but not to shy away from them either. Complex variations on training tasks and different camera orientations can be applied later in training, when trainees have attained a higher proficiency level. In this way, trainees are not overly challenged, but flexibility of their skills can still be enhanced.

Acknowledgements

We would like to acknowledge Gertjan Hultzer and René Rodenburg for their assistance in the LUMC Skillslab.

CHAPTER 07

INCREASING EFFICIENCY OF SURGICAL TRAINING: EFFECTS OF SPACING PRACTICE ON SKILL ACQUISITION AND RETENTION IN LAPAROSCOPY TRAINING

Edward N. Spruit MSc.
Dr. Guido P.H. Band
Prof. Jaap F. Hamming

Based on the publication:
Spruit, E. N., Band, G. P., & Hamming, J. F. (2015). Increasing Efficiency of Surgical Training: Effects of Spacing Practice on Skill Acquisition and Retention in Laparoscopy Training. *Surgical Endoscopy*, *29*(8), 2235-2243.

Abstract

Objectives: The goal of this study was to investigate the effects of spaced versus massed practice on skill acquisition and retention in the context of laparoscopic motor skill training.

Background: Reaching proficiency in performing laparoscopic surgery involves extensive training to acquire the required motor skills. Conventionally, training of such skills occurs during a full day training event utilizing surgical simulators that train specific motor skills pertinent to laparoscopic surgery. An important variable to consider is the optimal schedule for laparoscopic motor training.

Methods: In this study, two groups of trainees without prior experience were trained on a variety of physical box trainer tasks on different time-schedules. One group received three 75-minute training sessions on a single day (massed condition) and the other received one 75-minute training session per week for three consecutive weeks (spaced condition). Short- and long-term retention were assessed two weeks and one year after training completion.

Results: Outcome measures indicated better performance at the end of training, at a two-week delayed retention session and at a one-year retention session for the group that received training on a spaced schedule. This spacing effect was most pronounced for the more difficult laparoscopic training tasks such as intra-corporeal suturing. On average, 21% of participants in the massed group and 65% in the spaced group reach proficiency by the end of training.

Conclusions: Spacing practice of laparoscopic motor skill training will facilitate skill acquisition, short-term and long-term retention and thus, a more efficient learning process for trainees. Though more challenging in terms of logistics, training courses in medical centers should distribute practice sessions over longer time intervals.

Introduction

Since the rapid implementation of minimally invasive procedures at the end of the last century, the paradigm for training laparoscopic procedures to surgical residents has moved from the operating room to dedicated skillslabs for training purposes (Gallagher et al., 2005). A pressing research question is how to design training in the most efficient manner possible, while ensuring excellent skill acquisition, long-term retention and transfer to the occupational setting. Recent studies in the field of cognitive and educational psychology indicate that substantial improvements in learning efficiency can be achieved by an appropriate selection of feedback, proficiency targets (Gauger et al., 2010; Stefanidis et al., 2007), and video tutorials.

At least as important as the selection of material for laparoscopy training is an optimized dosage of delivering the training. Retention of training effects and transfer from trained to non-trained domains depend on factors such as deliberate practice, part-task training, task variability and overlearning after reaching proficiency (Stefanidis & Heniford, 2009; Adams, 1987; Spruit, Band, Hamming & Ridderinkhof, 2013). Most important for the current study, it has been well-documented[7] that distributing practice over time (spacing) leads to superior learning for knowledge acquisition, as well as motor skill acquisition.

In medical centers and hospitals, staff training is conventionally scheduled on a full day course because this is most convenient for organizing purposes. However, planners and curriculum designers need to ponder whether or not the benefits in terms of convenient logistics are worth the potential sacrifice in terms of the quality of learning. If the rate of skill acquisition suffers and long-term retention is compromised during training on a tightly crammed schedule, it may be wise to consider alternative planning methods for training medical staff.

Observations regarding the benefits of spacing practice are very robust in memory tasks, but are also prevalent for motor learning (Adams, 1987; Rosenbaum, Carlson & Gilmore, 2001). In a meta-analysis on verbal memory tasks (Cepeda, Pashler, Vul, Wixted & Rohrer, 2006), it was demonstrated that the lag (time interval) in between training sessions should increase as a function of the retention interval, with an optimal lag at 15-20% of the time until the final test (McDaniel, 2012). Even though in several domains of motor skills the spacing effect is also reliably and consistently demonstrated (Shea, Lai, Black & Park, 2000; Lee & Genovese, 1988), estimations of its magnitude varies with the training context (Donovan & Radosevich, 1999) and task complexity. That is why the current study will show the value of spacing in laparoscopic training for basic and advanced procedures.

The spacing effect has recently been researched in the setting of surgical training courses. Moulton and colleagues (2006) demonstrated significantly better retention following training on a microvascular anastomosis course for a group that received four training sessions in subsequent weeks (spaced) as compared to all on the same day (massed). In a different study, the spacing effect was tested while teaching laparoscopic cutting using the Minimally Invasive Surgical Training Virtual Reality (MIST-VR) (Gallagher, Jordan-Black & O'Sullivan, 2012). Performance was better if three training sessions were scheduled on consecutive days as compared to all on a single day.

Most spacing studies use small time intervals (minutes, hours or days, instead of weeks, months or years) (Donovan & Radosevich, 1999) out of logistical convenience (Cepeda et al., 2006) for the same reasons that trainers usually opt for massed training; it's just more practical. However, using small spacing windows provides little empirical basis for real educational settings where short-term and long-term retention are more important. The current study aims to incorporate skill retention and to differentiate among different levels of task complexity.

In the current study we aimed to replicate the spacing effect in a physical box-trainer model using an array of different laparoscopic training tasks varying in difficulty, using a weekly time interval for the spaced training group and adding short- and long-term retention of two weeks and a year, respectively. We hypothesize that the spaced group will have superior performance both at the end of training and at the retention sessions.

Methods

Participants

Fourty-one medical students (25 female) without prior experience in laparoscopy training were enrolled in the study. Age ranged from 17-28 (mean = 20) and all participants were right-handed. Participants received a certificate upon completion of the training as a compensation for taking part in the study.

Apparatus

Participants received training on a laparoscopic box trainer including four basic and one advanced task, all with previously established construct validity (Kolkman, van de Put, Wolterbeek, Trimbos & Jansen, 2008). All of these tasks aim to train perceptual and motor skills such as depth perception, adapting to the fulcrum effect and instrument handling, all of which are essential to proficiency in laparoscopic surgery. The first task requires participants to stretch a rubber band around a set of 12 spikes. In this task a trainee learns to work with forces. In the second task, participants string a pipe cleaner through a set of four rings. This task aims to train bi-manual dexterity. The third task involves the placements of small beads on a pegboard and requires very astute precision of motor actions. In the fourth task, a circle is cut in a rubber glove, which trains participants in exposure and dissection skills. In the advanced task, participants trained the skill of intra-corporeal suturing. Participants were taught how to create three knots, starting with the needle in their right instrument, using two throws for the first knot. One throw was used for the second knot starting in the left instrument and one throw for the third knot starting from the right instrument (see online video appendix in the reference list for all our laparoscopic training tasks*). During training, an open model of suturing was utilized to prepare participants in suturing before practicing in the laparoscopic box trainer.

Performance of participants on the box-trainers was recorded on a connected PC by means of video splitter and grabster (Terratec Grabster AV 400 MX) to convert video output to USB. The USB signal was converted to separate .mpg files by VLC Media Player for Windows.

Participants filled out self-report questionnaires covering demographics (sex, age, etc.), prior sport, music and gaming experience (0 = no experience, 1 = I used to play, 2 = yearly, 3 = monthly, 4 = weekly, 5 = daily), goal orientation (Seijts & Latham, 2006) and growth mindset (Dweck, 2006).

Training Programs

Training was given to 21 participants on a spaced schedule, and 20 participants on a massed schedule. Participants were randomly assigned to the two groups. All participants spent one hour on a set of psychological tasks (testing cognitive flexibility and spatial skills) prior to laparoscopy training. These cognitive tasks were hypothesized to predict skill acquisition on laparoscopic skills, but are beyond the scope of the present article. Both groups received laparoscopy training for a total of 225 minutes. This total training time was divided into three blocks of 75 minutes, which consisted of 15 minutes of instructions and 60 minutes of hands-on practice. In the first block, participants trained on the four basic laparoscopic tasks. During the second and third block, participants trained on all five laparoscopic tasks and an open suturing model to learn the basics of suturing prior to intra-corporeal suturing. During each block, participants completed each task twice in a fixed order (rubber band, pipe cleaner, beads, circle, suturing), after which participants were allowed to spend any remaining time on any training task of choice.

For the massed practice group, the three blocks of training were scheduled consecutively on one day. For the spaced practice group, these three blocks were separated by one week. After two weeks, a short-term retention session was scheduled to assess the participants' skill without any prior practice during that session. A long-term retention session was planned 12-14 months after training. Participants did not train their laparoscopy skills outside of the allocated training time.

Performance was video-recorded at the end of the first and third block of training and at the start of both retention sessions, totaling four moments of measurement for the first four tasks and three for intra-corporeal suturing.

Participants received standardized instructions by the trainer and self-directed feedback (Strandbygaard et al., 2013) in order to minimize confounding effects on the learning curve of the trainees.

Outcome measures

The video files of the participants were assessed by the first author for completion times of the task, as well as accuracy. An accuracy scoring tool based on principles of metrics by Gallagher and O'Sullivan (2012) was created for each task: frequently occurring steps and errors were scored and summed to form an accuracy measure for each of the laparoscopic tasks. Lower scores on completion times and accuracy (lower number of steps and errors made to complete a task) reflect better performance.

Statistical Analysis

Data were checked for normality and statistical tests were chosen accordingly. We tested whether groups were comparable at baseline in terms of age, sex, hand preference, musical, gaming, sports activity and personality factors.

For each of the five tasks (both for completion times and accuracy), for all moments of measurement (training session I, training session III, short- and long-term retention session) Mann-Whitney tests were performed to check for differences between groups at each stage of training. Wilcoxon signed-rank tests were performed as well to verify the improvements within trainees between training sessions. Furthermore, non-parametric correlations were used to explore potential relationships between questionnaire variables and performance on the laparoscopic tasks.

Results

Three of forty-one participants did not fully complete the training and were excluded from analysis. 38 participants ($N_{Massed} = 18$, $N_{Spaced} = 20$) took part in the short-term retention session and 12 ($N_{Massed} = 5$, $N_{Spaced} = 7$) participants completed the long-term retention session.

If at certain points during measurement participants were unable to complete the task within a reasonable amount of time (maximum of ten minutes), a score of 601 (a score that would automatically be assigned as the highest rank in the non-parametric tests) was assigned in order to avoid selective drop-out from our sample based on poor performance. This was the case for 12 out of 592 moments of measurement. Also, 28 out of 592 video files were lost due to trouble with the video recording equipment.

Baseline Check

Chi-square tests showed no significant differences between the two groups in terms of sex and hand preference. Mann-Whitney tests indicated no difference for gaming and sports activity, but the spaced group practiced significantly more (Mdn = 2 out of 5) with musical instruments than the massed group (Mdn = 0 out of 5), U = 64, z = -3.549, p < .001, r_{rb} = -0.64. Also, the spaced group was significantly younger (Mdn = 18.5 versus Mdn = 21), U = 43, z = -4.066, p < .001, r_{rb} = -0.76.

Main analysis: Basic laparoscopic tasks

The results for the first four tasks are shown in Figures 1-2. They indicate improvements in performance for all participants and highlight the differences in learning curves between the spaced and massed training groups.

At baseline, participants in the two groups did not show significantly different levels of performance on the four basic tasks, with the exception of completion times and accuracy scores on the rubber band task, with scores in favor of the spaced group.

At the end of training, performance levels on each of the first four tasks showed significant effects in favor of the group of the spaced training schedule, with the only exception of accuracy scores on the cutting circle task. Estimates of effect sizes (r_{rb}) of training were 0.67, 0.73, 0.65 and 0.36 for completion times on the elastic band, pipe cleaner, beads and circle cutting task, respectively. Effect sizes of training for accuracy scores at the end of training for each task were 0.63, 0.57, 0.48 and 0.13, respectively.

At the two-week post-training retention session, some of the differences in skill level on the first four tasks were still present, while others had vanished (see Figure 1-2). Effect sizes for the retention session for completion times on each task were 0.31, 0.42, 0.45 and 0.29, respectively. Accuracy scores effect sizes at retention were 0.16, 0.36, 0.31 and 0.11, respectively.

At one-year retention, the effects on the pipe cleaner task and the accuracy scores for the rubber band task persisted. Effect sizes were 0.07, 0.73, 0.17, 0.2 for completion times and 0.73, 1, 0.4, 0.5 for accuracy. The Wilcoxon signed-rank tests revealed the degree of within-person improvement between sessions and show a general pattern of improvements of trainees' performance from the first to the third training session for the first four tasks (see Figure 1-2).

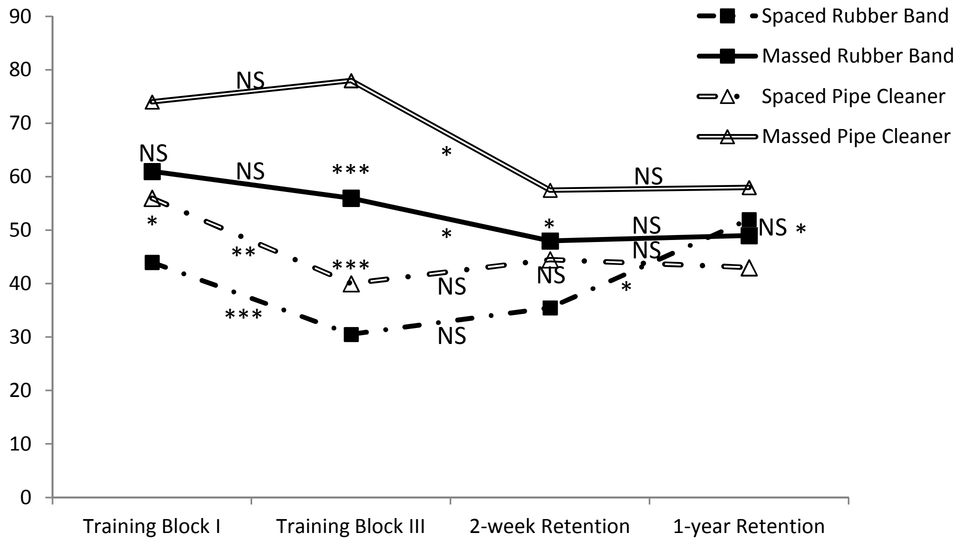


Figure 1a.

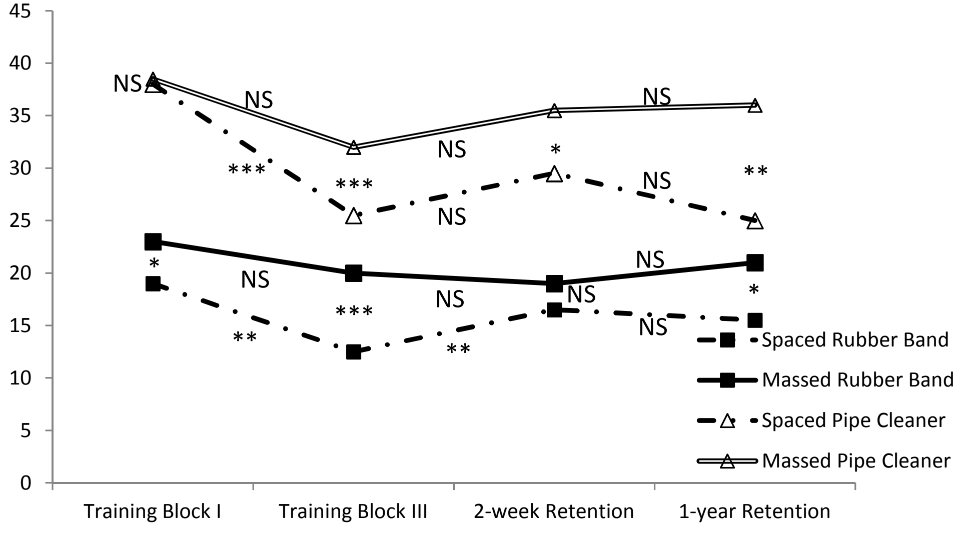


Figure 1b. Median completion times (1a) and median accuracy scores (1b) for the first two basic tasks after the first block of training (N=38), at the end of training (N=38) and at short-term (N=38) and long-term retention (N=12) for both training groups (NS = non-significant; * = p < .05; ** =p < .01; *** = p < .001).

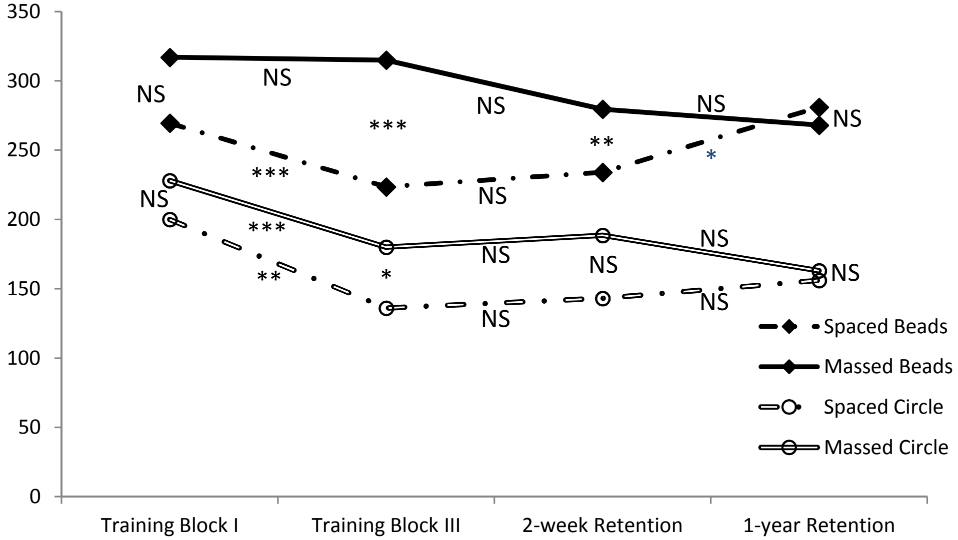


Figure 2a.

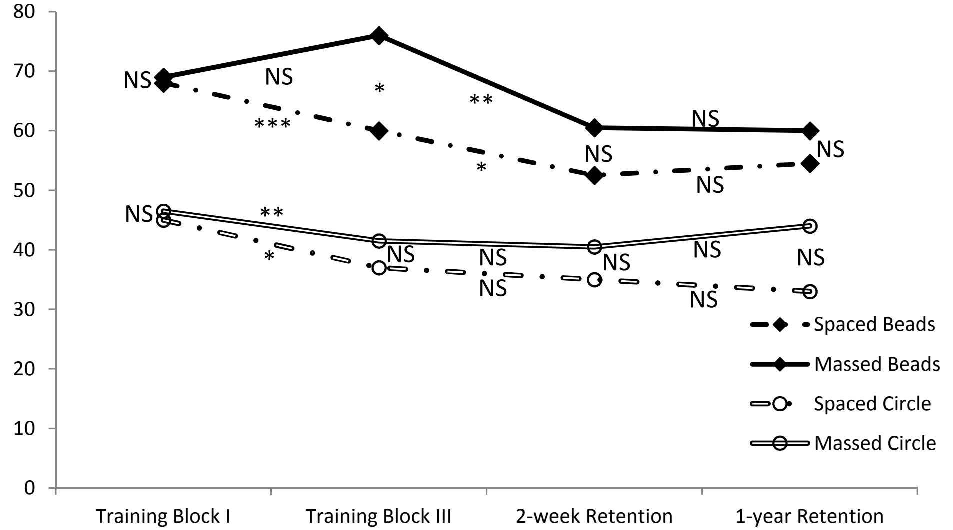


Figure 2b. Median completion times (1a) and median accuracy scores (1b) for the third and fourth basic task after the first block of training (N=38), at the end of training (N=38) and at short-term (N=38) and long-term retention (N=12) for both training groups (NS = non-significant; * = p < .05; ** =p < .01; *** = p < .001).

Main analysis: Intra-corporeal suturing

The Wilcoxon signed-rank tests for completion times and accuracy of intra-corporeal suturing showed no significant progress between the measurement at the end of training and short-term retention, as well as long-term retention (see Table 1-2).

Table 1. Mann-Whitney and Wilcoxon signed-rank tests to compare median completion times (in seconds) for the advanced task (intra-corporeal suturing)

	Mdn_{Massed}	Mdn_{Spaced}	U	Z	p
Training Block III	**325**	**180**	**71.5**	**-3.006**	**.001**
Short-term Retention	**271.5**	**183.5**	**84.5**	**-2.793**	**.002**
Long-term Retention	**488**	**275.5**	**4**	**-2.196**	**.014**
	NS	**NS**			

Table 2. Mann-Whitney and Wilcoxon signed-rank tests to compare median accuracy scores for the advanced task (intra-corporeal suturing)

	Mdn_{Massed}	Mdn_{Spaced}	U	Z	p
Training Block III	**74**	**42.5**	**77**	**-2.837**	**.002**
Short-term Retention	**60**	**42**	**87.5**	**-2.706**	**.003**
Long-term Retention	**113**	**39**	**0**	**-2.739**	**.002**
	NS	**NS**			

Mann-Whitney tests revealed substantial effects of group on completion times and accuracy scores on the intra-corporeal suturing task, both at the end of training and at the two retention sessions. Estimates of effect sizes (r_{rb}) for completion times at the end of training, short- and long-term retention were 0.58, 0.53 and 0.77, respectively. Effect sizes for accuracy scores were 0.55, 0.51 and 1, respectively. These results of intra-corporeal suturing at the end of training and the retention sessions for both groups are illustrated in Figure 3.

Extended analysis and confound check

Completion times on the different laparoscopic tasks correlated moderately with each other, with non-parametric Spearman's rho varying between .055 (ns) and .733 (p < .01) for completion times and from r = -.208 (ns) to r = .764 (p < .01) for accuracy. Correlations among completion times and their corresponding accuracy measures were very high, varying between r = .582 (p < .01) and r = .929 (p < .01), which indicates that accuracy on any given moment of measurement is highly related to completion times on that particular instance of performing a laparoscopic task and that participants did not trade accuracy for speed.

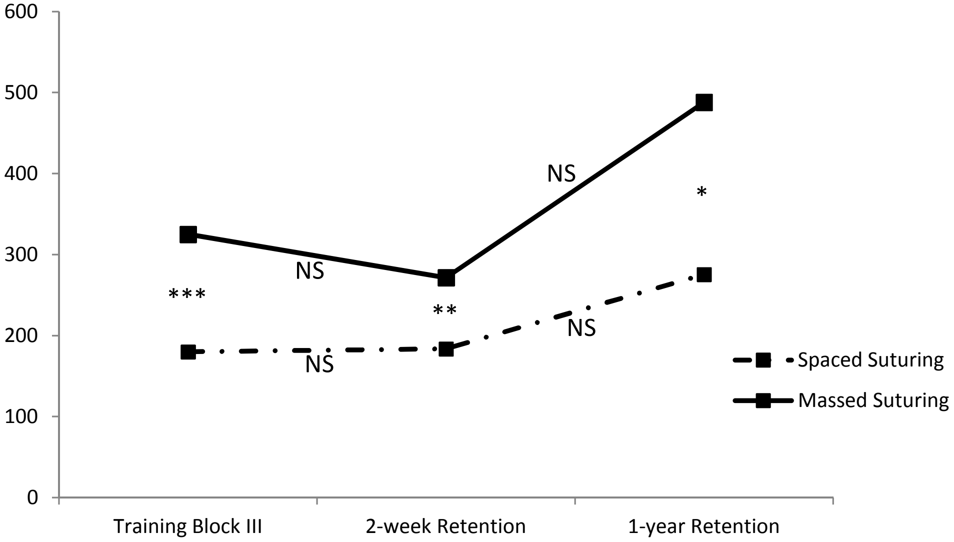


Figure 3a.

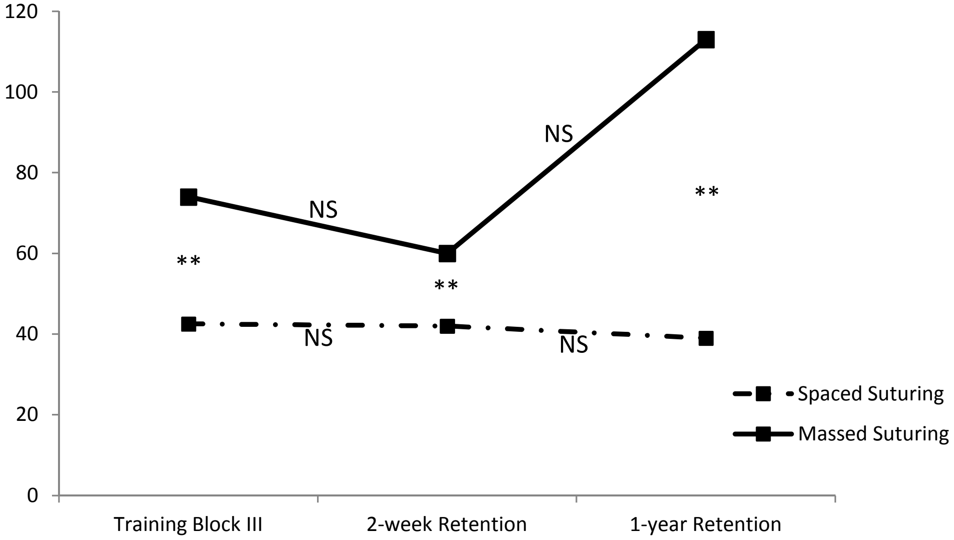


Figure 3b. Median completion times (1a) and median accuracy scores (1b) for the advanced task after the first block of training (N=38), at the end of training (N=38) and at short-term (N=38) and long-term retention (N=12) for both training groups (NS = non-significant; * = p < .05; ** =p < .01; *** = p < .001).

Further analysis showed no significant relations between sex, gaming activity, goal orientation, growth mindset and performance on any of the laparoscopic tasks. Significant correlations were found between age and some of the laparoscopic tasks on some of the moments of measurement, varying in magnitude from .338 to .636. Similarly, some correlations were significant for musical activity the laparoscopic tasks, ranging from -.351 to -.519.

To test the possibility that the factors of age and musical activity confounded the effects of spacing, we did a post-hoc case-controlled analysis for both variables. After matching both groups in age by gradually excluding the youngest participants from the spaced group and the oldest participants from the massed group until the groups were comparable in age, we found no major changes relative to the results of our main statistical tests. Matching groups for musical activity also did not substantially alter our results.

Cost-benefit analysis

To assess the success of our training, we compared both groups to a previously established performance benchmark for each task[16] (that is also used to determine whether a trainee has reached proficiency and is qualified to perform minor laparoscopic surgery in the OR). For the elastic band task, 7 out of 18 participants (39%) in the massed group have reached the proficiency benchmark by the end of training. For the spaced group, this number is 18 out of 20 participants (90%). For the pipe cleaner task, 2/18 (11%) versus 14/20 (70%) participants have reached proficiency. For the beads task, this comparison is 2/18 (11%) versus 10/20 (50%). For the cutting circle task: 4/18 (22%) versus 12/20 (60%). For intra-corporeal suturing: 4/18 (22%) versus 11/20 (55%).

Discussion

The current study replicated and extended previous studies showing that laparoscopic skills can be acquired in less training time by presenting a spaced schedule rather than the more typical massed schedule (Gallagher, Jordan-Black & O'Sullivan, 2012). After the same time investment, a larger proportion of students met proficiency criteria in the spaced than in the massed condition. Their performance was higher, clearly illustrated in lower completion times and accuracy scores. Moreover, we showed superior short- and long-term retention of the advanced suturing task up to a year after spaced as compared to massed training. Thus, the spaced schedule helps to maintain long-term reliability of skills, which evidently has implications for patient safety and training efficiency.

Overall, the spacing benefits were most pronounced for advanced skills, although benefits were also demonstrated for most of the indices of the basic skills. The relatively strong spacing effect for advanced skills is counter to what other motor skill research (Donovan &

Radosevich, 1999) suggests, since an earlier meta-analysis showed that the spacing effect usually diminishes with increasing task complexity.

This highlights the importance of scientific testing of learning strategies in unique training contexts. This finding clearly illustrates that trainees require less training time on a spaced schedule, which means less resources will be spent on training surgical residents.

The results showed minimal differences in the two groups in terms of demographics and initial performance. Therefore, the groups can be classified as comparable and the differences in performance and learning rates later in training can be attributed to our manipulation in the training set-up. Overall, the differences in groups on the first four tasks are not as pronounced as for intra-corporeal suturing (see Figure 3). It could be that participants needed less time to master the more basic tasks, resulting in a less pronounced difference in end levels of performance between the two groups after 225 minutes of training. Intra-corporeal suturing is a more cognitively demanding task and it typically takes much more practice to reach proficiency on it. Hence, task difficulty may have a moderating influence on the degree of the spacing effect in laparoscopy training.

It is also interesting to observe that in certain cases, participants in the massed group showed improvements in performance from the end of training to the retention session. For example, in accuracy scores on the beads/pegboard task and completion times for the rubber band and pipe cleaner task. In between these moments of measurement, there was no additional practice. By probing retention, there were two weeks of spacing built in for all participants, which may be a plausible explanation for this improvement.

A key question is what processes explain the benefits of spacing over massing training. The spacing effect can be explained in several ways. Obviously, trainees become mentally fatigued (van der Linden, 2011) after prolonged training. Fatigue has been found to impair learning of psychomotor and cognitive skills in laparoscopic tasks (Kahol et al., 2008). Thus, spacing training across multiple sessions can be beneficial by preventing fatigue.

A second explanation can be found in a differential effort investment. Every time a trainee starts a new training session, there is a gap to get performance back up to par (to the proficiency target). This gap is typically smaller on massed training sessions, where the knowledge and practiced skills remain active in working memory throughout the session with little effort investment by the trainee. Trainees on a spaced practice schedule do not have this advantage, as training information needs to be reactivated at the start of each session. This forces the trainee to exert more effort to attain the proficiency goal, which facilitates skill acquisition.

Furthermore, massed schedules lead trainees to overestimate how well they have mastered the skills in the training (Bjork, 1999), as it is easier to reproduce the same level of performance after a short time interval (same day) as compared to a longer time interval (a week later). Hence, massing training is beneficial for performance during training in the short-term at the risk of inaccurate appraisal of the trainees' actual skill level (Bjork, 1999). This incorrect form of self-efficacy poses a threat to the appropriate assessment of proficiency by both the trainee and the trainer. When proficiency is determined directly after a massed training, it gives an inaccurate picture of a trainee's skill level at the transfer setting (i.e. the first laparoscopic procedure for the trainee taking place several weeks/ months from now). It is therefore important to assess proficiency both after training and at a retention interval in order to ensure accurate skill assessment.

In training, most learning takes place in between practice, rather than during practice. Memory for motor skills (Donovan & Radosevich, 1999) improves due to consolidation, the gradual strengthening of memory that takes place in the elapsing time window that follows practice, to a large extent during sleep (Stickgold, 2005). When training of one motor skill is directly followed by training of a second motor skill, learning of the first skill is substantially impaired, a phenomenon known as retrograde interference (Brashers-Krug, Shadmehr & Bizzi, 1996). The explanation for this is that the new synaptic patterns (acquired during training) in the motor memory regions of the brain did not have any opportunity to process and consolidate and get overwritten by a new motor pattern during training of the second skill. This impairment vanishes when more time (four hours or more) elapses between training of the first and second skill. The positive effects of consolidation accumulate in this time window, but also during overnight sleep. Retrograde interference can be partially mitigated by a nap in between training of different motor patterns (Korman et al., 2007). This finding highlights the important role sleeps plays in the amount of consolidation that will take place in memory.

In other settings involving motor skill, such as dancing, one of the main advantages of spacing practice is that it reduces overuse injuries and improves recovery after training (Batson, 2007). This applies to laparoscopic surgical training as well, since many of our participants complained about minor pain in their hands and wrists after practicing the tasks for a prolonged time. This has mostly to do with the fact that they are novices and have a non-optimal posture, but spacing training immediately alleviates this problem.

All of these processes (mental fatigue, investment of effort in learning, accuracy of self-efficacy appraisal and memory consolidation) influence the advantages in learning that spacing offers, but it is unclear to what degree each of these adds to the effect of spacing. A more elaborate design would be needed to separate for example the influence of consolidation during sleep from just recovering from fatigue. In future studies, the time in between training sessions could be varied in order to further optimize training and nuance whether the advantages of spacing are mostly the effect of the draining of cognitive resources after a certain time on training or that consolidation of learning plays a more predominant part. Also, in this sample we initially kept a small short-term retention interval due to logistic convenience (low drop-out by participants), and a larger sample for long-term retention would be desirable in the future. It is noteworthy however, that in spite of the large participant drop-out on long-term retention, the effect for intra-corporeal suturing remained prevalent.

On a methodological note, we found that accuracy scores show a very similar pattern as completion times, which makes logical sense. Improvements in proficiency seem to follow a similar pattern in terms of completion times and accuracy. The accuracy measure as an assessment tool in differentiating experimental groups may seem redundant due to the high correlation with completion times. However, this accuracy assessment can complement the information derived from completion times to provide specific feedback for improvements while coaching trainees. Some trainees show a more reserved and conservative approach while performing a task, whereas other participants are less patient and make more steps and errors per unit of time while completing a task. The former would be more suited as a surgical trainee. Hence, accuracy scores are useful for individual assessment of proficiency during selection and examination.

A drawback in this study is that we originally intended to measure performance of the basic and advanced tasks at the end of training block II. Unfortunately, the majority of participants in the spaced group were unable to complete the intra-corporeal suturing task by training block II, which meant the allotted time for measurement would be exceeded if all tasks were to be recorded. In order to keep the length of training blocks equal in both conditions and stay on schedule, measurements for training block II were discarded.

Also, we used physical box trainers in a skillslab setting, which limits the extent to which our findings can be generalized to laparoscopy training in the OR. Participants acquired basic laparoscopic motor skills in our training, which does not take into account other important skills required for performance in the OR (such as navigation skills, decision making, team dynamics and knowledge of anatomy, patient and procedure).

The effect of spacing does not only have patient safety implications; there are also financial advantages. For a training institute a spaced schedule requires fewer resources (lab reservations, laparoscopic simulators, mentoring staff) for training surgical residents. Additionally, spacing different types of learning activities can enhance trainees' engagement in their training programs (Loehr & Schwartz, 2003). Since the advantages of spacing proved to be substantial, we recommend trainers to implement spaced practice in their surgical training curriculum. Scheduling training will perhaps be somewhat less convenient in terms of logistics, but the benefits in the quality of learning will outweigh the effort.

Acknowledgements

We would like to acknowledge Gertjan Hultzer, Rene Rodenburg and Tom Hulscher for their assistance in the LUMC skillslab.

CHAPTER 08

THE EFFECTS OF SPACING, NAPS AND FATIGUE ON THE ACQUISITION AND RETENTION OF LAPAROSCOPIC SKILLS

Edward N. Spruit MSc.
Dr. Guido P. H. Band
Dr. Kristiaan B. van der Heijden
Prof. Jaap F. Hamming

Based on the publication:
Spruit, E. N., Band, G. P., van der Heijden, K. B., & Hamming, J. F. (2017). The Effects of Spacing, Naps, and Fatigue on the Acquisition and Retention of Laparoscopic Skills. *Journal of Surgical Education, 74*(3), 530-538.

Abstract

Background: Earlier research has shown that laparoscopic skills are trained more efficiently on a spaced schedule compared to a massed schedule. The aim of the study was to estimate to what extent the spacing interval, naps and fatigue influenced the effectiveness of spacing laparoscopy training.

Methods: Four groups of trainees (aged 17-41 years; 72% female; $N_{massed} = 40$; $N_{break} = 35$; $N_{break\text{-}nap} = 37$; $N_{spaced} = 37$) without prior experience were trained in three laparoscopic tasks using a physical box trainer with different scheduling interventions. The first (massed) group received three 100-minute training sessions consecutively on a single day. The second (break) group received the sessions interrupted with two 45-minute breaks. The third (break-nap) group had the same schedule as the second group, but had two 35-minute powernap intervals during the breaks. The fourth (spaced) group had the three sessions on three consecutive days. A retention session was organized approximately three months after training.

Results: The results showed an overall pattern of superior performance at the end of training and at retention for the spaced group, followed by the break-nap, break and massed group, respectively. The spaced and break-nap group significantly outperformed the break and massed group, with effect sizes ranging from .20 to .37.

Conclusions: Spacing laparoscopic training over three consecutive days or weeks is superior to massed training, even if the massed training contains breaks. Breaks with sleep opportunity (i.e. lying , inactive, muted sensory input) enhance performance over training with regular breaks and traditional massed training. For optimal skill acquisition and retention of laparoscopic skills, larger spacing intervals (at least up to a week) are recommended.

Introduction

Acquiring laparoscopic motor skills requires an extensive amount of practice. Earlier research suggested that practice time allocated for surgical training can be used most efficiently when scheduled across multiple smaller time intervals (Moulton et al., 2006; Gallagher, Jordan-Black & O'Sullivan, 2012), preferably with several non-training days in between training sessions. In an earlier study (Spruit, Band & Hamming, 2014, chapter 7), we established that laparoscopic skills are better acquired and retained when learned on a spaced schedule as compared to a massed schedule. In the study, two groups of participants learned four basic and one advanced task (intra-corporeal suturing) on a physical box trainer. The first group received training on a massed 1-day schedule, while the second group received the same amount of training divided over three sessions spaced across three consecutive weeks. Performance at the end of training, at a two week retention session and a one year retention session was superior for the spaced group. These spacing effects were most pronounced for the advanced task.

Spacing training provides multiple benefits. Having multiple shorter training sessions reduces the likelihood of trainees becoming overly fatigued (Kahol et al., 2008) or bored, which may be the case when practicing laparoscopic tasks for an extended amount of time. Also, trainers and trainees get a more accurate reflection of their actual skill level during spaced training (Bjork, 1999). Since the training uses multiple sessions, trainees will discover they do not perform as smoothly at the start of the second training session as they did at the end of the first training session. This gives them a more accurate appraisal of their own skill level. The fresh start at the beginning of each new training session requires trainees to invest more effort to get their performance back up to the level they left at the end of the previous session. This allows for more elaborate and frequent activations of the neural pathways in the brain associated with learning the specific motor skill.

Theorists have suggested that consolidation of memory is adaptive (Wang, Zhou & Shah, 2014), and that learning a skill is favored on a spaced training schedule because frequent short practice sessions on a training task give the brain the indication that it will encounter the same task more often in the future. This triggers more enduring memory encoding in order to accommodate for repetitive encounters with the task. Spacing provides time intervals between training sessions during which memory consolidation (Stickgold, 2005) can be strengthened. This in turn supports longer retention. Consolidation occurs in the brain when a person is disengaged from the trained activity and is enhanced during sleep (Brashers-Krug, Shadmehr & Bizzi, 1996), when perceptual input to the brain is reduced to a large extent. In designing training, one ought to be cognizant of the fact that most learning takes place during off-line periods, not during practice. During training, practice provides the learning input that will be consolidated at a later period of time post-training. If too many similar training tasks are practiced right after another, the learning input created by the first training experience will be replaced and thus impaired by later training experiences (Stickgold, 2005). A more adequate strategy is to provide trainees rest intervals so the associated brain regions can have the opportunity to process the learning input before proceeding with more training. Research has shown consolidation is enhanced by overnight sleep (Walker, Brakefield, Morgan, Hobson & Stickgold, 2002) as well as power naps (Korman et al., 2007).

From our earlier results (Spruit, Band & Hamming, 2014, chapter 7), we were unable to distinguish to what extent each of these factors facilitates acquisition and retention of laparoscopic motor skills acquired on a physical box trainer. In the current study we use a more elaborate design in order to dissect the individual effects that might contribute to spacing training.

The goal of this study was two-fold: (1) to compare the effectiveness of an intervention with a smaller spacing interval to a bigger one (one day versus one week) and (2) to determine the influence of factors such as naps, mental fatigue and simple breaks on the skill acquisition and retention of laparoscopy training. With the current study we aimed to attain valid estimates of the extent to which each of these factors influences the effectiveness of spacing training.

First, we hypothesize that groups that have a larger time interval between training sessions will have superior performance at the end of training and at retention and experience lower levels of fatigue. Second, we hypothesize that sleep opportunities in between training sessions will lead to superior performance at the end of training and at retention and result in lower levels of fatigue.

Methods

Participants

149 university students (108 female) without any prior experience in laparoscopy training were enrolled in the study. Age ranged from 17-41y (mean = 21) and 128 participants were right-handed. All subjects filled out informed consent forms and were granted a training certificate as a reward for participating in the study.

Apparatus

Participants trained three laparoscopic tasks on a physical box trainer. All tasks have previously established construct validity (Kolkman, van de Put, Wolterbeek, Trimbod & Jansen , 2008). The tasks train perceptual and motor skills such as depth perception, adapting to the fulcrum effect and instrument handling, all key skills for mastering laparoscopic surgery. In the first task, participants had to coordinate a pipe cleaner through a set of four rings. This task is utilized to improve a trainees bi-manual dexterity. In the second task, participants had to pick up small beads from a bucket and drop them on pins on a pegboard making the shape of a simple figure. This task requires caution and very careful handling of the pins with the instruments. In the final task, participants learn intra-corporeal suturing. The suturing sequence required correct insertion of the needle and three knots. The first knot required two throws, while the second and third knot required one throw. Participants were instructed to start with the needle in their right instrument for the first and third knot and to start with the needle in their left instrument for the second knot. This was done to unsure flexibility of the skill for both hands (see online video appendix listed in the reference list for all the laparoscopic training tasks[*]). During training, participants also learned suturing on an open model in order to prepare them before practicing in the laparoscopic box trainer.

Video footage of the performance of participants on the box-trainers was converted to .mpg files for each task at each moment of measurement using a video splitter, grabster (Terratec Grabster AV 400 MX) and VLC Media Player for Windows.

Participants filled out self-report questionnaires covering demographics (sex, age, etc.), prior sport, music and gaming experience (0 = no experience, 1 = I used to play, 2 = yearly, 3 = monthly, 4 = weekly, 5 = daily), personality (Gosling, Rentfrow & Swann, 2003) and mental fatigue (RSME, Rating Scale Mental Effort, Zijlstra, 1993).

Training Programs

Training was divided into three sessions. The duration of each session was 100 minutes, divided into 50 minutes of practice on the training tasks, 15 minutes of instructions and 35 minutes of measurement (laparoscopic tasks, questionnaires). All three laparoscopic tasks were practiced and measured during each of the three sessions, but the duration of practice for the pipe cleaner and beads task was reduced for the later sessions, while practice time on intra-corporeal suturing was increased. Total practice time for the pipe cleaner, beads and intra-corporeal suturing task was 27.5, 32.5 and 60 minutes, respectively. During the first session, measurement took place before practice to establish a baseline. At the baseline measurement for intra-corporeal suturing, only the insertion of the needle in the model is performed, since the knot tying part of the task is too difficult for a novice without any prior training. For the second and third session, measurement took place after practice at the end of the session. RSME questionnaires were filled out right before measuring the laparoscopic tasks on each session. A retention session was scheduled approximately three months after training. All three laparoscopic tasks were measured during retention without any prior practice.

Participants received standardized instructions by the trainer and instructional videos (see online appendix in reference list*) and no feedback in order to minimize confounding effects on the learning curve of the trainees. If trainees asked for feedback they were reminded to pay close attention next time the instructional video would be shown, since all the required information to learn is present in the video.

There were four groups, each with a different time schedule. The spaced group received the training sessions on three consecutive days. The massed group had all sessions consecutively on a single day without any breaks. The break and break-nap group received the sessions on one day with two 45 minute breaks in between, with the break-nap group having a 35 minute powernap opportunity during both breaks. In the break-nap group, participants had a powernap opportunity on inflatable mattresses in a dark room while wearing earplugs and a sleeping mask to minimize sensory input. After each powernap, a self-report sleep questionnaire was filled out. All participants in the break and break-nap group wore activity trackers (Flex, Fitbit) on their wrist to measure their activity levels during training and during the breaks. The instrument uses three-dimensional motion sensing technology, records data in one-minute epochs, and has shown to be an acceptable instrument for sleep/wake monitoring in normative populations (Montgomery-Downs, Salvatore & Bond, 2012). Participants were assigned randomly to each group. All groups trained their laparoscopic skills for an equal amount of time.

Performance outcome measures

The video files of the participants were assessed by the first author for completion times of the task. If participants were unable to complete the task within a set amount of measurement time (maximum of ten minutes during training and fifteen minutes during retention), a score of 601 or 901 (seconds) was assigned in order to avoid selective drop-out from our sample based on poor performance. This score would automatically be assigned as the highest rank in the non-parametric tests. Thus, there were three outcome measures (pipe-cleaner, beads, suturing task) at each moment of measurement (at baseline, end of session 2, at the end of training, and retention).

A total score was computed as a fourth measure using z-scores from the laparoscopic tasks (we chose this method in order to make sure all three tasks contribute equally to the total score, since the laparoscopic tasks have differing mean completion times).

Statistical Analysis

Data were checked for normality and statistical tests were chosen accordingly (One-way ANOVA, independent samples and paired samples t-tests in the case of normal data and Kruskal-Wallis, Mann-Whitney and Wilcoxon Signed-rank tests in the case of non-normal data) using the statistical software SPSS 23.0. A significance level of .05 was used. We tested whether groups were comparable at baseline in terms of age, sex, hand preference, academic year, food, caffeine, alcohol intake, sleep prior to training, mental fatigue, musical, gaming, sports activity, and personality factors. For our main analysis, we tested all four laparoscopic scores (pipe cleaner, beads, suturing and total scores) on all four moments of measurement (baseline, at the end of session two, at the end of training and at the retention session) with 16 Kruskal-Wallis tests.

Results

149 participants (N_{massed} = 40; N_{break} = 35; $N_{break\text{-}nap}$ = 37; N_{spaced} = 37) completed the training and 134 participants (N_{massed} = 36; N_{break} = 27; $N_{break\text{-}nap}$ = 35; N_{spaced} = 36) returned to the lab for the follow-up retention session.

Participant characteristics

Chi-square tests showed no significant differences between all the groups in terms of sex, hand preference and food intake. Caffeine intake was lower in the break-nap group (35.1%) compared to the other groups (massed: 62.5%; break: 51.4%; spaced: 70.2%; p = .022). Alcohol intake (on the night prior to training) differed between groups (massed: 62.5%; break: 37.1%; break-nap: 35.1%; spaced: 13.5%; p < .001). In order to check for any effects of caffeine or alcohol intake on performance on the laparoscopic tasks, eight Mann-Whitney tests were performed (two within each of the four groups). None of the tests were significant, indicating no influence of caffeine or alcohol intake in the sample.

Mann-Whitney tests revealed no significant differences between groups in age, musical gaming and sports activity. However, the break-nap group had more participants in the initial years of their study ($\text{Median(Mdn)}_{break\text{-}nap}$ = 1.0y; Mdn_{massed} = 3.0y; Mdn_{break} = 3.0y; Mdn_{spaced} = 2.0y).Non-parametric correlations showed no significant relations between years of study and performance on any of the laparoscopic tasks.

One-way ANOVAs showed no significant differences between groups in any of the big five personality factors and hours of sleep (the night prior to training). For quality of sleep (the night prior to training) the results were mostly comparable, with the exception of a slightly lower sleep quality in the break group compared to the break-nap group (p = 0.03, $Mean_{break}$ = 3.38, SD_{break} = 0.92; $Mean_{break\text{-}nap}$ = 3.92, $SD_{break\text{-}nap}$ = 0.76), but without differences for the other two groups ($Mean_{massed}$ = 3.78, SD_{massed} = 0.66; $Mean_{spaced}$ = 3.57, SD_{spaced} = 0.83). Also, baseline mental fatigue was significantly higher in the spaced group compared to the break-nap group (p = 0.001, $Mean_{spaced}$ = 22.53, SD_{spaced} = 12.25; $Mean_{break\text{-}nap}$ = 12.42, $SD_{break\text{-}nap}$ = 8.40), but without differences for the other two groups ($Mean_{massed}$ = 18.94, SD_{massed} = 12.10; $Mean_{break}$ = 21.45, SD_{break} = 12.18). Non-parametric correlations showed no significant relations between sleep quality (the night prior to training) or baseline mental fatigue and performance on any of the laparoscopic tasks.

Results of the activity trackers showed a large difference in the estimated minutes spent asleep between the break and the break-nap group for both of the two breaks (Break 1: $p < 0.001$; $\text{Mdn}_{\text{break}} = 0$; $\text{Mdn}_{\text{break-nap}} = 10$; Break 2: $p < 0.001$; $\text{Mdn}_{\text{break}} = 0$; $\text{Mdn}_{\text{break-nap}} = 12$). Self-reported minutes asleep in the break-nap group ranged from 0 to 30 ($\text{Mean}_{\text{nap1}} = 5.95$; $\text{Std.Dev.}_{\text{nap1}} = 6.62$; $\text{Mean}_{\text{nap2}} = 6.73$; $\text{Std.Dev.}_{\text{nap2}} = 8.22$). The activity trackers' measure of minutes spent asleep correlated significantly with the associated self-report measure ($r = .353$, $p = .032$). No significant correlations were found between self-reported sleep the night prior to training and self-reported (r = .-178, $p = .291$) and activity trackers' measure ($r = -.148$, $p = .218$) of minutes spent asleep during the breaks, although the coefficients were in the direction one would expect.

Table 1. Mean ranks (from multiple Kruskal-Wallis Tests) for all four training groups at baseline, the end of session II, at the end of training ($N_{\text{massed}} = 40$; $N_{\text{break}} = 35$; $N_{\text{break-nap}} = 37$; $N_{\text{spaced}} = 37$) and at retention ($N_{\text{massed}} = 36$; $N_{\text{break}} = 27$; $N_{\text{break-nap}} = 35$; $N_{\text{spaced}} = 36$).

	Mean Ranks			
	Massed	**Break**	**Break-nap**	**Spaced**
Measure:				
Pipe cleaner$_{\text{baseline}}$	76.95	71.37	69.47	81.85
Pipe cleaner$_{\text{end of session II}}$	78.95	79.03	62.36	77.66
Pipe cleaner$_{\text{end of training}}$	84.71	81.37	65.58	67.89
Pipe cleaner$_{\text{retention}}$	82.99	79.56	55.70	54.44
Beads$_{\text{baseline}}$	82.61	70.21	73.78	70.49
Beads$_{\text{end of session II}}$	82.85	72.36	75.82	64.07
Beads$_{\text{end of training}}$	84.22	72.89	69.09	71.19
Beads$_{\text{retention}}$	74.97	74.63	63.71	56.42
Suturing$_{\text{baseline}}$	71.41	76.79	76.42	73.81
Suturing$_{\text{end of session II}}$	82.98	79.39	72.03	61.46
Suturing$_{\text{end of training}}$	86.78	72.94	72.15	63.32
Suturing$_{\text{retention}}$	81.10	73.00	57.37	59.63
Total z-scores$_{\text{baseline}}$	77.63	72.37	73.29	72.30
Totsl z-scores$_{\text{end of session II}}$	83.25	76.58	69.00	66.70
Total z-scores$_{\text{end of training}}$	88.51	77.06	64.50	65.05
Total z-scores$_{\text{retention}}$	83.42	79.74	54.71	52.64

Main analysis: Laparoscopic tasks

For our main analysis, the mean ranks of the Kruskal-Wallis tests are shown in Table 1. If differences between groups were observed, an independent samples Mann-Whitney test was performed to compare individual groups with each other. Figures 1-4 show the results of these tests with asterisks to flag whenever the outcome of these Mann-Whitney tests proved significant. Figures 1-3 show median scores with 25 to 75 % inter-quartile ratio's on each task at each session and serve as an indication of the different learning curves between the training groups. Note that the baseline of the suturing task is lower since only a smaller segment of the task was performed at baseline. This measure was only used for baseline between group analysis, not within-person progress.

At baseline, none of the groups showed significantly different levels of performance on any of laparoscopic tasks. At the end of session two, we observed a significant difference on the pipe cleaner task between the break and the break-nap group ($p = 0.044$), while the spaced group performed better at the beads task ($p = 0.024$), the suturing task ($p = 0.012$) and the total score ($p = 0.044$) as compared to the massed group. The spaced group also performed significantly better than the break group ($p = 0.047$) on suturing at the end of the second session. At the end of training both the spaced and break-nap group outperformed the massed group on the pipe cleaner ($p = 0.044$; $p = 0.034$, respectively) and the total score ($p = 0.007$; $p = 0.009$, respectively).

The spaced and break-nap group also showed a lower score on the suturing task compared to the massed group, but only borders close to significance for the break-nap group ($p = 0.008$; $p = 0.056$, respectively). At retention, both the spaced and break-nap condition outperformed the massed ($p = 0.001$; 0.001) and the break group ($p = 0.005$; $p = 0.01$, respectively) on the pipe cleaner task. On the beads task, the spaced group significantly outperformed both massed ($p = 0.017$) and break ($p = 0.032$) groups. On the suturing task, both the spaced and break-nap group outperform the massed group ($p = 0.007$; $p = 0.005$, respectively). Both the spaced and break-nap group have a significantly lower total score compared to the massed ($p < 0.001$; $p = 0.001$, respectively) and the break group ($p = 0.002$; $p = 0.007$, respectively).

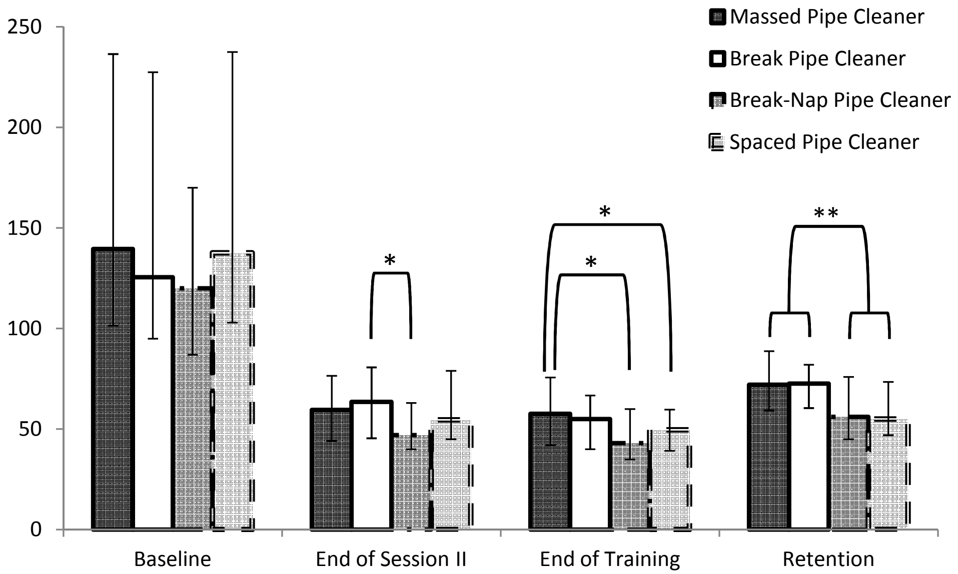


Figure 1. Median completion times (in seconds) for the pipe cleaner task at baseline, at the end of the second session, at the end of training and at retention for all training groups (* = p < .05; ** = p < .01; *** = p < .001).

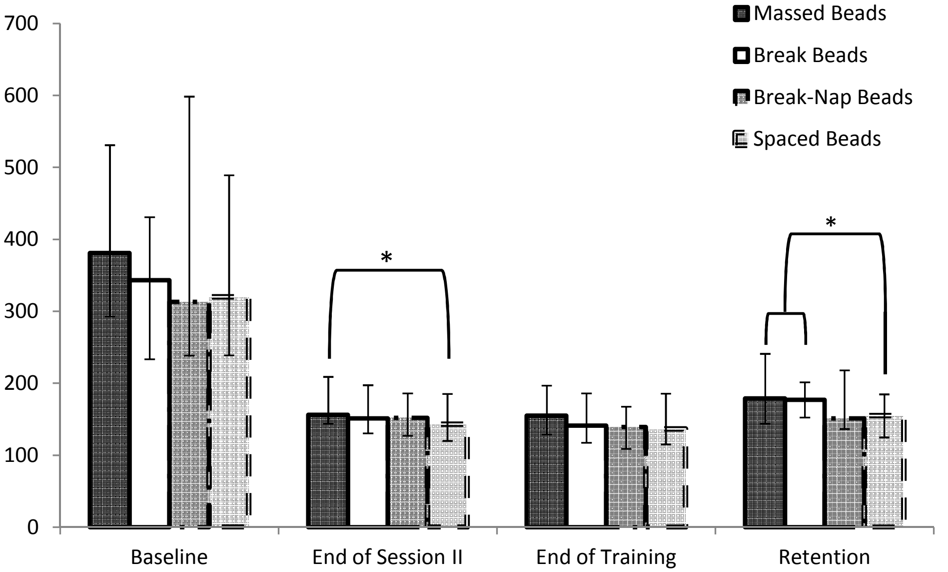


Figure 2. Median completion times (in seconds) for the beads task at baseline, at the end of the second session, at the end of training and at retention for all training groups (* = p < .05; ** =p < .01; *** = p < .001).

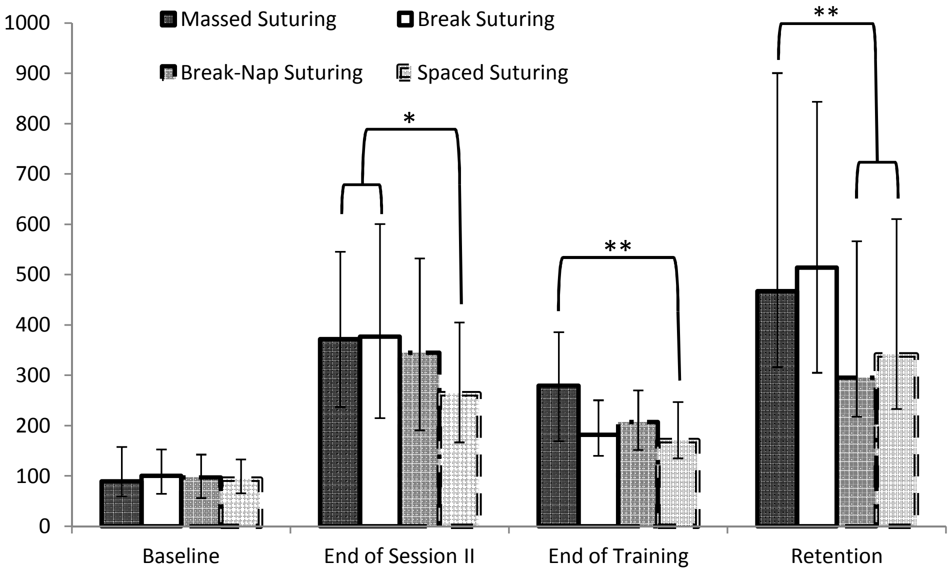


Figure 3. Median completion times (in seconds) for the intra-corporeal suturing task at baseline, at the end of the second session, at the end of training and at retention for all training groups (* = p < .05; ** =p < .01; *** = p < .001). Please note that only a segment of the suturing task was performed at baseline (see method section), hence the lower completion times.

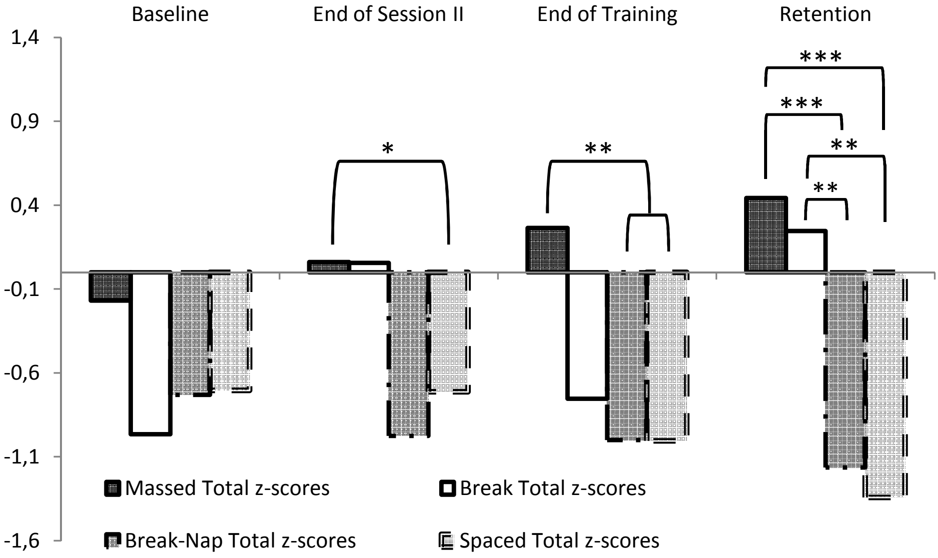


Figure 4. Median total z-scores at baseline, at the end of the second session, at the end of training and at retention for all training groups (NS = non-significant; * = p < .05; ** =p < .01; *** = p < .001).

One week versus one day spacing comparison

In sum, the Mann-Whitney tests revealed substantial differences in completion times on the laparoscopic tasks between the spaced and massed group. Estimates of effect sizes (r_{rb}) for completion times at the end of the second session, at the end of training, and at retention are displayed in Table 2, along with the effect sizes found in a prior study that used a spacing interval of one week. Since the retention interval was different from the current study, only end of training comparisons have been made.

Sleep and laparoscopy measures

To differentiate between the effects of estimated sleep and mere rest (and muted sensory input), we assessed the relation between self-reported sleep at breaks, self-reported sleep the night prior to training, the results of the activity trackers at breaks, and performance on the laparoscopic tasks within the subsample of the break-nap group.

Non-parametric correlations showed coefficients ranging from r = .040 (p = 0.813) to r = .465 (p = 0.004) for self-reported sleep during the naps and performance on the subsequent laparoscopic tasks. These correlations were most prevalent for the end of training and retention measures, but also present at baseline. However, none of the laparoscopy measures were significantly correlated with the results of the activity trackers' estimate of minutes spent asleep. No significant correlations were found between performance on the laparoscopic tasks and self-reported sleep the night prior to training.

Table 2. Effect sizes (r_{rb}) of the Mann-Whitney tests for the three training tasks. Effect sizes are only mentioned when tests were significant and when a comparison was viable (for the previous spacing study).

	break-nap|break	break-nap|massed	spacing|break	spacing|massed	spacing|massed Spruit et al. (2014)
Pipe cleaner task					
End of Session II	0.20				
End of Training		0.21		0.19	0.73
Retention	0.29	0.35	0.32	0.37	
Beads task					
End of Session II				0.23	
End of Training					0.65
Retention			0.23	0.25	
Suturing Task					
End of Session II			0.20	0.26	
End of Training				0.27	0.58
Retention		0.31		0.29	

Table 3. Means Rating Scale Mental Effort scores with paired samples t-tests assessing the progression in mental fatigue for all four training groups at baseline, the end of session II and at the end of training ($^{\approx}$ for no significant difference; $^{\downarrow}$/ $^{\uparrow}$ for significant increase/ decrease, * = p < .05; ** = p < .01; *** = p < .001).

	Mean$_{\text{massed}}$	**Mean**$_{\text{break}}$	**Mean**$_{\text{break-nap}}$	**Mean**$_{\text{spaced}}$
Measure:				
RSME$_{\text{baseline}}$	18.94	21.45	12.42	22.53
RSME$_{\text{end of session II}}$	33.81 $^{\uparrow***}$	40.67 $^{\uparrow***}$	35.43 $^{\uparrow***}$	25.15 $^{\approx}$
RSME$_{\text{end of training}}$	50.61 $^{\uparrow***}$	48.50 $^{\approx}$	48.30 $^{\uparrow***}$	27.55 $^{\approx}$

Rating Scale Mental Effort Analysis

No significant correlations were found between scores on the RSME and estimated time spent asleep at breaks (self-report and activity tracker). At baseline and at the end of session two, no effects of mental fatigue on laparoscopic performance were found. Fatigue at the end of session three did show significant correlations with the pipe cleaner (r = .192, p = 0.019), suturing (r = .206, p = 0.012) and total score (r = .219, p = 0.008) at the end of training and the suturing task at retention (r = .173, p = 0.046).

The RSME scores are shown in Table 3., along with the results of multiple paired samples t-tests illustrating the progression of the scores for all four training groups. A clear reduction in mental fatigue can be observed for the spaced group compared to the three other groups.

Discussion

The findings in this study further nuanced the current theory regarding the spacing effect as it applies to laparoscopic surgical training. We found that spacing three training sessions across three consecutive days is advantageous compared to a massed 1-day schedule, even if that schedule accounts for substantial amounts of breaks in training sessions.

Alternatively, we found that including powernaps in between the training sessions on a 1-day schedule, can enhance long-term retention, an option which may be more beneficial in terms of logistics when organizing training events, when time constraints are a bigger motive.

The Effects of Naps

One of the more striking findings was that within the break-nap group, participants who reported spending more time asleep during the naps, had worse performance on the laparoscopic tasks at the end of training and retention. This seems counter-intuitive since the break-nap group as a whole shows a pattern of better performance compared to the break group, the most similar condition that does not allow sleep during the break.

A possible explanation is that the lower performance was caused by sleep inertia, a time period of impaired performance and disorientation as a person transitions from sleep to wakefulness (Miccoli, Versace, Koterle, Cavallero, 2008). We accounted for this possibility in the schedule of our sessions, since all performance measurements were done directly prior to the naps, with the naps being followed by a brief period to fill out a questionnaire and resuming training on a basic task in the next session. Still, sleep inertia may have effected trainees while practicing during the following session or the inertia effects may not have fully wore off by the time of the next measurement (Burke, Scheer, Ronda, Czeisler & Wright, 2015). The occurrence of sleep inertia can differ between contexts. The recommendation on the duration for daytime naps is to keep them below 30 minutes, as longer naps are associated with lower productivity and increased sleep inertia (Dhand & Sohal, 2006). However, when employees work in night shifts or have prolonged working hours, longer naps (40-60 minutes) show more performance benefits (Mulrine, Signal, Berg & Gander, 2012). Additionally, effects on motor performance can be influenced by whether a trainee habitually takes naps or not (Milner, Fogel & Cote, 2006).

Another explanation is that participants who slept (longer) within the break-nap group did so because they were more sleep deprived and sleepy to begin with, and this led to a suboptimal state for learning throughout the day of training (Miccoli et al., 2008), while the naps were not sufficient to decrease sleep deprivation and sleepiness. Contrary to this explanation, there was no association between sleep quantity and sleep quality on the night before the training and performance on the laparoscopic tasks (both within the break-nap group and the overall sample).

An alternative explanation is that within the break-nap group, trainees with a lower aptitude for learning laparoscopic skills experienced more arousal due to their difficulties in learning the tasks and reported being more fatigued and having spent more time asleep during the breaks as a misattribution of their arousal levels (Cotton, 1981). In any case, the findings should be interpreted cautiously since no correlations were found between the performance on the laparoscopic tasks and estimated sleep derived from the activity monitoring devices. Regardless of whether sleep during naps deteriorates performance, the results do indicate that trainees benefit from a period of rest while they are inactive, lying down, and while their sensory input is muted as compared to a traditional break.

The Effects of Fatigue and Consolidation

Even though participants in the massed group reported experiencing significantly more fatigue compared to the spaced group, we observed weak correlations between fatigue and performance on the laparoscopic tasks in the sample, which suggests a small influence of fatigue contributing to the spacing effect. It is also worth considering that fatigue may have a higher impact on learning efficiency when a higher level of fatigue is reached, but that this threshold was simply never reached using the current design of training.

Learning of motor skills largely occurs in the timeframe following training (Walker, Brakefield, Morgan, Hobson & Stickgold, 2002), when there is opportunity for consolidation of the training stimuli. Consolidation occurs during 'off-line' periods simply as time passes, but is enhanced with a short nap or during overnight sleep (Brashers-Krug, Shadmehr & Bizzi, 1996; Doyon et al., 2009; Korman, Raz, Flash & Karni, 2003), although this may depend on the nature of the task (Doyon et al., 2009). Typically, effects of overnight sleep consolidation are enhanced for more difficult tasks (Kuriyama, Stickgold & Walker, 2004) and when a break is planned early in the training schedule rather than later (Duke, Allen, Cash & Simmons, 2009). Other authors have suggested that sleep effects can be attributed to the type of design used (sleep deprivation control groups that tend to impair performance) and that mere periods of rest show similar improvements as sleep (Rieth, Cai, McDevitt & Mednick, 2010). Regardless of whether sleep enhances consolidation or whether sleep deprivation impairs performance (or both), one can conclude that sleeping sufficiently is beneficial after training. The results of the current study also suggest that night sleep is more beneficial in between training sessions than just after the entire training, since the spaced group (who had two nights of sleep in between training sessions) had the best performance at the end of training and at retention, followed by break-nap group, the break group and the massed group, respectively.

When comparing the two spaced groups (the one in the current study and the one from the previous study, Spruit, Band & Hamming, 2014, chapter 7), we observed higher effect sizes for the intervention with the bigger spacing interval (one week). Future research should investigate what the optimal spacing interval is, although this is likely influenced by the desired retention interval (McDaniel, 2012). Finally, trainers ought to be cautious in the design of their training, since consolidation can also be disrupted during the acquisition (active training) phase when multiple conflicting training conditions are interleaved during training (Banai, Ortiz, Oppenheimer & Wright, 2010; Spruit, Kleijweg, Band & Hamming, 2016, chapter 6).

Limitations and Conclusion

One of the limitations of the study is that we did not include any measures of self-efficacy in our design. In the introduction we noted that spacing can enhance more accurate appraisal of the skill level of a trainee (by both trainee and instructors) and with the current design we were unable to estimate the extent to which this factor influenced the effectiveness of spacing. Also, we used a different retention interval from the previous study (Spruit, Band & Hamming, 2014, chapter 7) so we only have comparisons of effects sizes for the end of training and not retention. The current study only used completion times, so no generalizations to other performance outcome measures (accuracy, instrument path length, force, etc.) can be made from these findings.

Also, we used self-reports and activity monitoring devices to measure sleep, but these only serve as an approximation of time spent asleep. Future studies could make use of polysomnography to provide a more valid estimate on the influence of sleep consolidation on skill acquisition and retention of laparoscopic motor skills.

To answer the initial research question, we conclude that consolidation has the biggest influence in the effectiveness of spacing. Mitigating mental fatigue by inclusion of breaks and providing small periods of sensory muting are beneficial, but play a smaller role. When applying powernaps in training, one ought to be cautious for sleep inertia and it is recommended to schedule short (below 30 minutes, Dhand & Sohal, 2006) with ample rest time after waking to allow for waning of potential sleep inertia. Another way of preventing a poor learning state and poor consolidation post-training is to instruct trainees to sleep sufficiently before and after each training session. Finally, we found that spacing over three weeks leads to better learning efficiency and retention than spacing over three days.

Acknowledgements

We would like to acknowledge Gertjan Hultzer, Rene Rodenburg, Racheal Cheung, Joram van Ketel, Jolanda Hus, Tirza Loman, Remco Duijn, Enola van Maarsseveen and Frederique Arntz for their assistance in the LUMC skillslab.

CHAPTER 09

REFLECTIONS, GENERAL DISCUSSION AND RECOMMENDATIONS

Partially based on the publication:
Spruit, E. N., Band, G. P., Hamming, J. F., & Ridderinkhof, K. R. (2014).
Optimal Training Design for Procedural Motor Skills: a Review and Application
to Laparoscopic Surgery. *Psychological Research, 78*(6), 878-891.

Reflections, General Discussion and Recommendations

In the current project, our main focus was to test the effectiveness of different training interventions and their impact on skill acquisition and long-term retention of laparoscopic motor skills. In the literature, there is a substantial amount of prospective studies that investigate the influence of different trainee factors on the individual aptitude to acquire laparoscopic skills.

Training Interventions

In the third chapter, we listed several interventions which may be applied in laparoscopy training. Based on an extensive survey of the relevant literature, we submit several recommendations for optimal training design for procedural motor skills, and LS in particular.

1) Surgical training models need to be validated in terms of construct and predictive validity. An expert benchmark of performance needs to be established for the different training tasks in the simulator. The markers for performance can vary among the type of model that is used, but in general we advocate objective measures that are most crucial to surgical performance. This means that we value clearly defined metrics, completion times, accuracy and force feedback over path length and subjective rating scales. The expert benchmark serves as a training goal of proficiency for each specific motor skill for trainees. Although this benchmark is determined in terms of performance, the focus in designing training should be on learning.
2) Trainees should engage in adaptive training on a spaced practice schedule. This will allow for trainees to progress in the required motor skills at their individualized pace and for optimal allocation of cognitive resources for skill acquisition and long-term retention.
3) Segmented part-task training can be applied in complex training tasks to reduce the cognitive load and facilitate learning, although more research is needed to determine the validity of this learning strategy.
4) Dual-task training conditions should be implemented to measure the degree of automatization of motor skills, as well as to 5) provide adequate integration of multiple tasks present in the criterion task. These immersive training conditions should closely resemble the realistic transfer setting of the OR.
6) Care should be taken that there is variability in task parameters during training, as this is associated with better long-term retention of skill.
7) Training sessions ought to be spaced in order to reduce mental fatigue during training and to increase consolidation, resulting in enhanced skill acquisition and retention.

8) Mental imagery can be utilized in surgical skill acquisition to reduce the amount of time and materials required for training.
9) The curriculum should capitalize on deliberate practice of skills that require further development. By continuing to engage executive control in learning after proficiency has been reached, trainees remain engaged in the process of continual improvement that will help them make the transition to expert level in the years following their training.

In addition to these points, which were discussed in detail in the above sections, a number of additional recommendations for medical training made by Wulf et al. (2010) deserve to be cited here. Wulf et al. recommended self-controlled practice, which corresponds with learner control (Wickens, Hutchins, Carolan & Cumming, 2011) in that trainees have an influence in deciding which task they practice on. Such a strategy is consistent with the aim to adapt the training to individual competence levels, but as indicated before, trainees are not necessarily good at deciding which skill is mastered and which requires additional training.

In addition, Wulf et al. (2010) argued in favor of observational practice by dyads, which is a cost-effective way of training in which two trainees can learn from each other in pairs. Providing feedback during training can have informational, as well as motivational effects on a trainee. We emphasize the importance of self-directed feedback during the acquisition of complex LS skills. Typically, minimal instructor feedback is needed during basic tasks of low difficulty, but trainees can benefit from self-directed feedback during more difficult tasks (Strandbygaard et al., 2013). Continual feedback may help reduce cognitive load during LS tasks, but this reduction in load does not necessarily reduce the intrinsic load of a task. If an instructor tells a trainee exactly what to do during a task this is more likely to take away from germane load, which makes performing the task too easy for the trainee. This type of continual feedback will make it less likely that a trainee is actively engaged in learning the task and can lead to inaccurate estimates of competence of a trained skill (Bjork, 1999).

Integration with Empirical Chapters

Into many of these interventions, empirical studies were conducted during this research project. The use of proficiency targets we did not assess with an individual study in the current dissertation, since this training principle has been researched thoroughly in the existing literature and is evidently beneficial for trainees. We did provide all trainees in our studies with existing proficiency targets for each task and developed the training from there. In the fifth chapter, we reaffirm that the use of visual force feedback to the trainee can be used effectively to train participants in bringing down the amount of force they exert on the training tissues. This study is unique in this dissertation in that it utilizes additional different measures besides task completion times and accuracy scores. From the data in this study, we found that force parameters reflect a different construct than the often correlated completion times and instrument path length. This highlights the use of force as an important metric for safe tissue handling skills in laparoscopy training, next to the pre existing metrics that signify efficiency.

In the sixth chapter, we explored the feasibility of a varied practice intervention. The results of this study revealed a negative effect for the group with more variability of practice. The pattern of results suggests this was mostly the result of using inverse viewing conditions on the laparoscopic rather than the effect of frequent interleaving of training tasks. In the seventh and eighth chapter we investigated the effects of spacing practice. These interventions showed the biggest benefit for skill acquisition and retention of laparoscopic skills out of all the empirical studies that were performed during this research project. From the second spacing study, we conclude that one week of spacing is beneficial over one day. This suggests that consolidation from mere spacing (and overnight sleep) plays a bigger role in facilitating learning than the alleviating of fatigue by rest.

Finally, there are a number of studies which did not add up to a full chapter, but the results of which will be briefly mentioned here. One study tested a part-task condition (which segmented training of intra-corporeal suturing in seven smaller sub-tasks) and a mental imagery condition against a control condition. The sub-tasks consisted of: (1) correct positioning of the needle in the instrument and correct orientation of the instrument into the laparoscopic simulator; (2) inserting the needle into the foam model with the right instrument, grabbing the other end of the needle with the left instrument completing the insertion and creating the correct length of the suture with the right instrument (while keeping vision on the needle) to start knot tying; (3) switching the needle from the left to the right instrument grabbing it by the tip and achieving correct orientation of the needle in the right instrument; (4) making two loops (throws) of suture around the left instrument by moving it in a circular motion around the end of the needle where the suture is attached; (5) grabbing the end of the suture with the left instrument, letting the loops slide off slowly by pulling the right instrument (which still holds the needle); (6) switching the needle over to the left instrument in the correct orientation, making one loop (throw), grabbing the end of the suture and completing the second knot; (7) switching the needle over to the right instrument in the correct orientation, making one loop (throw), grabbing the end of the suture with the left instrument and completing the third knot (see the online video appendix in the reference list to get a visual representation of the task). The part-task group practiced each of the sub-task separately and only performed that portion of the task during that part of the training. At the later stages of training, trainees performed the whole task together.

Participants in the mental imagery condition received a hidden internet link with the instructional video's used during training and were instructed to watch these at home in between the weekly training sessions and to mentally rehearse themselves performing all the tasks successfully, at least ten times per week.

A minor non-significant trend was observed in favor of the part-task training condition at the end of training compared to the other two conditions, but no differences were found between the mental imagery condition and control condition. It is worth noting that simulator training in its own right is already a form of part-task training, namely fractionation. Most laparoscopic simulators only train motor and perceptual skills and the training tasks are not completely comparable to surgical procedures. It is important to realize that training on a simulator that integrates more facets (perceptual, knowledge, social, technical) of laparoscopic surgery is desirable before heading on to the training in the operating room. Simulators that integrate more facets are starting to emerge, but many improvements can still be made.

In another study, we tested a set of gaming principles during laparoscopic training in order to facilitate more coaching and interpersonal communication between trainees. The experimental group were instructed that their training group was a team and it was their goal to achieve the best performance as a group, rather than an individual. Trainees were also incentivized with a prize which would be awarded to the best performing group. Prizes were awarded randomly in a control group who received the traditional training. Results indicated no significant differences between the two groups in acquisition of laparoscopic skills, but did reveal a higher rate of proactive social behavior in the experimental group. I do recommend this intervention since the application of this principle requires very little investment and benefits interaction between trainees without compromising training efficiency.

Predictors of Laparoscopic Skills

In each of the empirical studies performed, we measured a variety of background variables of the trainees. These variables were primarily assessed to ensure comparability between the experimental and control groups, but had a secondary purpose in that their relationship with laparoscopic skill acquisition and retention could be assessed. Over the multiple studies performed, overall results can be briefly reported. We did not find any significant correlations between sex, age, academic year, openness to experience, extraversion, conscientiousness, agreeableness, growth mindset, goal orientation, proactive social behavior, gaming and sports activity. The lack of an effect of age may be due to range restriction (most trainees were between the age of 17 and 28, predominantly in their early twenties). Overall, musical activity was correlated with performance on the laparoscopic tasks. It is not clear whether this relationship is causal (practicing a musical instrument improves laparoscopic skills) or shares an underlying construct that influences both proficiency with musical instruments and in laparoscopic skills.

Recommendations

This dissertation started with an analysis of laparoscopic surgery as a complex task performed by a surgeon in the operating room. This task entails many facets, one of which was the main focus of the remainder of this dissertation, namely the motor and partially perceptual aspects of training surgical residents in laparoscopy. The main focus for the current research project was laparoscopic motor skill training, but it is important to note that other facets are equally (or perhaps more) important when developing a curriculum for surgical residents. I suggest instructors to also explore the associated literature for these remaining facets. Also, I'd like to encourage researchers to design their future studies in different facets of laparoscopy training. Specifically, I'm referring to the other important skills a surgical resident ought to develop (i.e. surgical knowledge, perceptual skills, clinical decision making, mental endurance, social skills and technical skills, see chapter 2). This would especially be valuable on the topics where research is yet scarce in this relatively young field.

Based on the research in this dissertation I recommend instructors to design training with predetermined proficiency targets on a spaced schedule with intervals of a week instead of smaller time frames. Instructors may experiment with larger spacing intervals, but more research is needed to determine the effectiveness of more time in between training sessions.

I urge instructors to be cautious in increasing training variability in training novices, since laparoscopy is already an inherently complex task and can be overwhelming at the start of training. Fractionation of training of the different facets of laparoscopic surgery may be fine initially, but training focused on skill integration is desirable at a later stage. In examination, a dual-task setup can be used to assess the degree of automatization of the acquired skills. I encourage trainers to adopt force measures and visual force feedback as an additional metric in laparoscopy training to ensure trainees not only have fluent and efficient motor skills when they finish training, but also learn how to handle different tissues safely. Different outcome metrics of laparoscopic performance (completion times and exerted force) are not highly correlated and both efficiency and safe tissue handling are essential for being an excellent laparoscopic surgeon.

CHAPTER 10

A. SUMMARY

B. SAMENVATTING

10a. Summary

This dissertation describes research investigating ways of improving efficiency of laparoscopic motor skill training. Performing laparoscopic surgery is a complex task that involves using a small camera (laparoscope), a visual monitor and a set of long instruments to perform surgery in the abdomen. The laparoscope and instruments are inserted via small incisions to make the procedure less invasive and create smaller scars compared to traditional open surgery (laparotomy).

In the **Introduction (chapter 1)**, I describe how the increase of minimally invasive procedure such as laparoscopy has led to the demand for an alternative approach to surgical education. Traditionally, surgery is taught in a mentor-apprenticeship model, where a surgical resident first observes and assists an experienced surgeon and gradually starts to perform more challenging tasks in the operating room. Due to the complexity of laparoscopic surgery , this was no longer a viable option. This led to an increase in the development of surgical simulators, where a trainee can practice difficult surgical skills in a safe environment.

It also led to a stronger emphasis on standardized evidence-based medical education. In **chapter 2**, the results of a cognitive task analysis is reported, specifically describing all the essential action steps and decision points of an advanced laparoscopic procedure (resection of the sigmoid colon in the case of a tumor). It also lists what skills and knowledge ought to be included in a curriculum for laparoscopic residents with an estimate of their importance by a panel of experienced laparoscopic surgeons.

In **chapter 3**, the focus of the dissertation narrows down to simulation training of laparoscopic motor and perceptual skills. In this chapter, relevant literature is reviewed to provide a set of nine recommendations that trainers can use to enhance laparoscopy training. It suggests trainers use training models that have construct and preferably predictive validity with benchmarked proficiency levels by expert surgeons, that can serve as goals for trainees. Trainees should engage in adaptive training on a spaced practice schedule. Mental imagery can be utilized in surgical skill acquisition to reduce the amount of time and materials required for training. Trainers can use part-task training to reduce cognitive load, although care should be taken, since there are not a lot of studies that have tested this concept in medical training. Also, it is important to provide for a form of skill integration at a later stage in training. Dual-task training conditions can be used to measure the degree of automatization of motor skills, and to provide adequate integration of all the different tasks present in the operating room. The curriculum should actively endorse an attitude of deliberate practice and life-long learning, so residents continue towards surgical mastery after they have reached proficiency at the end of training.

Trainers can increase the variability of different task parameters at a later stage of training, as this is commonly found to correlate with better long-term retention of skill. In **chapter 6**, we found that increased variability of the angle of the laparoscope in the simulator hampered skill acquisition and retention in novice trainees, so we recommended lower amounts of task variability for novices. Differences in training task switching was not associated with worse or better performance.

Chapter 4 comprises a short commentary on a study on self-directed feedback touching on the theme of feedback dependency. Feedback is helpful tool to increase a trainee's pace of learning and also has a motivational function. When feedback is omnipresent, a trainee's performance (or confidence) can become too dependent on it and this can become problematic if no one is aware of how of a trainee's performance is reliant on the presence of (instructor) feedback.

In **chapter 5**, the subject of feedback continues, namely visual force feedback. In previous studies, the presentation of visual force feedback on the simulator monitor facilitated the acquisition of safe tissue manipulation skills as compared to a control group without visual force feedback. In the current chapter, the aim was to find the optimal dosage for the presentation of visual force feedback. The different feedback conditions consisted of: (1) continuous, (2) gradually fading in as time-on-task accumulates or (3) only when a trainee exceeds a certain force threshold. All conditions significantly reduced the amount of exerted force on the tissues, but no significant differences were found between groups.

In **chapter 7**, we compared a massed training group with a spaced training group. The massed group received all their laparoscopy training on a single day, whereas the spaced group had three training seasons spread across three consecutive weeks. Participants on the spaced group performed significantly better on the laparoscopic tasks at the end of training, two weeks after training and approximately a year after training. In **chapter 8**, the spacing effect was further studied with a larger sample and four different training groups. The first group was a massed condition, which has all of the training sessions on a single day without any major breaks in between. The second group also received training on a single day, but with two 45-minute breaks in between three training sessions. The third group was identical to the second, except that they were allowed two 30-minute powernap opportunities on an inflatable mattress in dark room with muted sensory input. The fourth group had three training sessions spaced over three consecutive days. In the results, we found that the third and fourth group outperformed the first and second group at the end of training and two months after training, but the margins and effect sizes were not as large (and not statistically significant for every task comparison) as the study described in **chapter 7**, in spite of larger sample sizes per group.

In **chapter 9**, the results of the training interventions tested in this dissertation are integrated with the recommendations from the literature review. I briefly mention the results of the prospective analyses (in regard to factors such as age, sex, sport, gaming, music activity, etc.) done on the data presented in this dissertation as well as two studies that did not make it as a full chapter into the final dissertation due to the lack of meaningful results and the absence of retention data.

10b. Samenvatting

Dit proefschrift beschrijft onderzoek gericht op het verkennen van verschillende manieren om de efficientie van laparoscopische motorische vaardigheidstraining te verbeteren. Het uitvoeren van laparoscopische chirurgie is een complexe taak, waarin gebruik gemaakt wordt van een kleine camera (een zogenaamde laparoscoop), een beeldmonitor en twee langwerpige instrumenten, waarmee chirurgische handelingen kunnen worden verricht in de buikholte. De laparoscoop en de instrumenten worden in de buikholte ingebracht via kleine incisies om de operatie minder invasief te maken en resulteren dan ook in kleinere littekens in vergelijking met traditionele open chirurgie (laparotomie).

In de **Introductie (hoofdstuk 1)** wordt beschreven hoe de toename van vormen van minimaal invasieve chirurgie (zoals laparoscopie) heeft geleidt tot een behoefte aan een alternatieve benadering voor chirurgische training. Van oudsher wordt chirurgie getraint in een mentor-leerling model, waarin de chirurg-in-opleiding allereerst de mentor (ervaren chirurg) observeert en daarna ook ondersteunt bij een laparoscopische operatie. In verloop van tijd verricht de leerling steeds uitdagendere en complexere handelingen in de operatie kamer. Omdat laparoscopische chirurgie veel complexere vaardigheden vereist, is dit model niet langer compleet houdbaar. Dit heeft ertoe geleid dat er meer chirurgische simulatoren ontwikkeld zijn, waarmee een leerling de complexe chirurgische vaardigheden in een veilige omgeving kan aanleren. Daarnaast is er hierdoor ook een sterkere nadruk gekomen op medisch onderwijs dat gebaseerd is op wetenschappelijke bevindingen.

In **hoofdstuk 2** worden de resultaten van een cognitieve taak analyse gerapporteerd. Hier worden de essentiele handelingen en beslismomenten van een geavanceerde laparoscopische procedure (laparoscopische dikke darm (sigmoid) resectie) beschreven. Dit hoofdstuk bevat daarnaast een lijst van belangrijke kennis en vaardigheden, welke men naar aanbeveling van ervaren laparoscopische experts zou kunnen opnemen in een curriculum voor chirurgen-in-opleiding.

In **hoofdstuk 3** vernauwd de focus van het proefschrift zich en specifeer ik het thema naar simulatie training in laparoscopische motorische and perceptuale vaardigheden. In dit hoofdstuk wordt een literatuurstudie beschreven, waaruit negen verschillende aanbevelingen volgen die trainers kunnen gebruiken om de aangeboden laparoscopische trainingen te verbeteren. Trainers worden geadviseerd trainingsmodellen/simulatoren te gebruiken die over een bewezen construct en (bij voorkeur) voorspellende validiteit beschikken. Daarnaast is het raadzaam, bekwaamheidscriteria (vastgesteld met behulp van laparoscopie experts) te hebben waar naartoe de leerlingen kunnen streven als leerdoel. De leerlingen kunnen adaptief getraind worden op basis van hun actuele vaardigheidsniveau op een tijdschema dat over meerdere trainingsdagen verspreid is.

Mentale verbeeldingsoefeningen kunnen gebruikt worden om de benodigde trainingstijd en -materiaal te verlagen. Trainers kunnen deel-taak training gebruiken om de cognitieve lading te verminderen. Voorzichtigheid wordt hierbij geboden, gezien het feit dat dit concept nog niet uitgebreid bestudeert is in medisch onderwijs. Tenslotte is het belangrijk dat er in de latere fasen van training een vorm van vaardigheidsintegratie aangeboden wordt. Dubbel-taak training condities kunnen gebruikt worden om de mate van automatisering van de motorische vaardigheden te meten. Daarnaast biedt het een goede gelegenheid om de verschillende taken te integreren die men in de operatiekamer uitvoert (motorische handelingen, relevante patient informatie, procedurele en anatomische kennis oproepen, met operatie kamer teamleden communiceren, informatie randapparatuur waarnemen). Het curriculum zou een houding van 'deliberate practice' en 'het leven lang blijven leren' moeten bijbrengen, zodat leerlingen ook nog na het afsluiten van hun training, hun hele carriere lang eraan blijven werken om betere chirurgen te worden.

In een later stadium van de training, kunnen mentoren de variabiliteit van verschillende taak parameters verhogen om de training complexer te maken. Dit kan het lange termijn behoud van aangeleerde vaardigheden verhogen. In **hoofstuk 6**, wordt een studie naar variabiliteit in laparoscopie training beschreven. Daar vonden we dat een hogere mate van variabiliteit in de ruimtelijke orientering van de laparoscoop leidde tot een lager leerniveau en verminderd behoud van de vaardigheden na de training bij beginnende trainees. Om deze reden, raden wij af om bij aanvang van de training veel variabiliteit aan te bieden. Frequenter van taak wisselen tijdens de training was niet geassocieerd met betere of slechtere prestatie.

Hoofdstuk 4 is een kort kommentaar op een studie over zelf-gedirigeerde feedback en gaat onder andere over het thema feedback afhankelijkheid. Feedback is een behulpzaam middel om het leertempo van de trainee te verhogen en heeft daarnaast nog een motiverende functie. Wanneer feedback alomaanwezig is kan dit ervoor zorgen dat een trainee's prestatie (of zelfvertrouwen) te afhankelijk wordt van de feedback. Dit kan voor problemen zorgen wanneer niemand zich tijdens het training proces bewust is van het feit dat de trainee's prestatie afhankelijk is van de aanwezigheid van feedback van een instructeur.

Hoofdstuk 5, gaat verder in op het onderwerp van feedback, met name visuele 'force feedback'. In voorgaande studies is gevonden dat de presentatie van visuele force feedback bevorderlijk werkt bij het aanleren van laparoscopische vaardigheden, met name gericht op het veilig manipuleren van weefsels. In de beschreven studie was het doel om te onderzoeken wat de optimale dosering van de presentatie van visuele force feedback is. De verschillende feedback condities waren: (1) continueel (doorgaand), (2) gradueel de presentatie van feedback verhogen gedurende de training of (3) alleen wanneer de trainee een bepaald krachtdrempel overschrijdt. Al deze drie condities zorgde voor een aanzienlijke afname in de kracht die trainees op de weefsels uitoefenden. Tussen de condities was geen significant verschil te vinden.

In **hoofdstuk 7**, worden twee groepen met elkaar vergeleken waarvan de ene groep een gespreide training en de andere groep een massa training volgden. De massa training groep volgde de gehele laparoscopie training op één dag, terwijl de gespreide training groep drie trainingsessies volgde op drie opeenvolgende weken. Deelnemers uit de gespreide groep presteerde significant beter op de laparoscopische taken aan het einde van de training, twee weken na de training en ongeveer een jaar na de training.

In **hoofdstuk 8**, wordt het thema van trainingspreiding verder bestudeerd met een grotere steekproef en vier verschillende traininggroepen. De eerste groep was een massa conditie, waarin de deelnemers alle training op één dag volgden, zonder grote pauzes tussen het trainen. De tweede groep volgde de training ook op één dag, maar met twee pauzes van 45 minuten tussen drie trainingsessies. De derde groep had een vergelijkbare tijdsplanning als de tweede groep, afgezien van het feit dat zij tweemaal de gelegenheid hadden om voor 30 minuten een dutje (powernap) te doen in een donkere ruimte (met oordoppen en slaapmasker). De vierde groep volgde drie trainingsessies verspreid over drie opeenvolgende dagen. Als resultaten, vonden we dat de derde en vierde groep beter presteerde dan de eerste en tweede groep aan het einde van de training en twee maanden na de training. De verschillen en effectgroottes waren, ondanks een grotere steekproef per groep, echter niet zo groot (en niet statistisch significant voor alle laparoscopische taken) als in de voorgaande studie beschreven in **hoofstuk 7**.

In **hoofdstuk 9**, worden de resultaten van alle in dit proefschrift geteste training interventies geintegreerd met de aanbevelingen uit de literatuurstudie. Daarnaast worden de resultaten van prospectieve data analyses met betrekking tot andere randfactoren (leeftijd, geslacht, sport, gaming, muziek activiteit, etc.) kort besproken. Ook is er een korte berichtgeving van twee uitgevoerde studies die uiteindelijk niet als een compleet hoofdstuk gepresenteerd konden worden, gegeven het gebrek aan betekenisvolle resultaten en gebrek aan retentie data.

CHAPTER 11

REFERENCE LIST

Reference list

*. Online appendix containing videos of our laparoscopic training tasks and also serve as an illustration of how our raw data was captured:
https://www.youtube.com/channel/UClH5gJxQ9lr9DXkUxtl7_9w

A

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological bulletin*, *101*(1), 41.

Aggarwal, R., & Darzi, A. (2006). Technical-skills training in the 21st century. *New England Journal of Medicine, 355*(25), 2695-2696.

Aggarwal, R., Moorthy, K., & Darzi, A. (2004). Laparoscopic skills training and assessment. *British Journal of Surgery, 91*(12), 1549-1558.

Ahlberg, G., Enochsson, L., Gallagher, A. G., Hedman, L., Hogman, C., McClusky, D. A., Ramel, S., Smith, C. D., & Arvidsson, D. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *The American journal of surgery*, *193*(6), 797-804.

Arora, S., Aggarwal, R., Sirimanna, P., Moran, A., Grantcharov, T., Kneebone, R., Sevdalis, N. & Darzi, A. (2011). Mental practice enhances surgical technical skills: a randomized controlled study. *Annals of surgery, 253*(2), 265-270.

ASGE training committee (2012). Small-bowel endoscopy core curriculum. Gastrointestinal endoscopy. doi:10.1016/j.gie.2012.09.023.

Avgerinos, D. V., Goodell, K. H., Waxberg, S., Cao, C. G. L., & Schwaitzberg, S. D. (2005). Comparison of the sensitivity of physical and virtual laparoscopic surgical training simulators to the user's level of experience. *Surgical Endoscopy And Other Interventional Techniques, 19*(9), 1211-1215.

Awad, S. S., Fagan, S. P., Bellows, C., Albo, D., Green-Rashad, B., De La Garza, M., & Berger, D. H. (2005). Bridging the communication gap in the operating room with medical team training. *The American Journal of Surgery*, *190*(5), 770-774.

B

Balch, C. M., Freischlag, J. A., & Shanafelt, T. D. (2009). Stress and burnout among surgeons: understanding and managing the syndrome and avoiding the adverse consequences. *Archives of surgery*, *144*(4), 371-376.

Banai, K., Ortiz, J. A., Oppenheimer, J. D., & Wright, B. A. (2010). Learning two things at once: differential constraints on the acquisition and consolidation of perceptual learning. *Neuroscience*, *165*(2), 436-444.

Banta, H. D. (1993). Minimally invasive surgery. Implications for hospitals, health workers, and patients. *British Medical Journal*, *307*(6918), 1546-1549.

Barreiros, J., Figueiredo, T., & Godinho, M. (2007). The contextual interference effect in applied settings. *European Physical Education Review*, *13*(2), 195-208.

Bashankaev, B., Baido, S., & Wexner, S. D. (2011). Review of available methods of simulation training to facilitate surgical education. *Surgical endoscopy, 25(1)*, 28-35.

Batson, G. (2007). Revisiting overuse injuries in dance in view of motor learning and somatic models of distributed practice. *Journal of Dance Medicine & Science*, *11*(3), 70-75.

Battig, W. F. (1966). Facilitation and interference. In E. A. Bilodeau (Ed.), *Acquisition of skill* (pp. 215−244). New York: Academic Press.

Bemelman, W. A. (2009). Minimaal Invasieve Chirurgie: Plan van aanpak en beleid. *Nederlandse Vereniging voor Endoscopische Chirurgie*.

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective?. *Cognitive psychology*, *61*(3), 228-247.

Bjork, R. A. (1999). *Assessing our own competence: Heuristics and illusions.* (in D. Gopher & A. Koriat (Eds.), Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application (pp. 435-459)).

Blavier, A., Gaudissart, Q., Cadière, G. B., & Nyssen, A. S. (2006). Impact of 2D and 3D vision on performance of novice subjects using da Vinci robotic system. *Acta Chirurgica Belgica*, *106*(6), 662-664.

Bourne, L. E., Jr. & Healy, A. F. (2012). *Introduction: Training and Its Cognitive Underpinnings.* (in A. F. Healy & L. E. Bourne Jr. (Eds.), Training cognition: Optimizing Efficiency, Durability, and Generalizability (pp. 1-12).

Boutin, A., & Blandin, Y. (2010). On the cognitive processes underlying contextual interference: Contributions of practice schedule, task similarity and amount of practice. *Human movement science, 29*(6), 910-920.

Brady, F. (2004). Contextual Interference: a Meta-analytic Study. *Perceptual and motor skills*, *99*(1), 116-126.

Brashers-Krug, T., Shadmehr, R., & Bizzi, E. (1996). Consolidation in human motor memory. *Nature, 382*(6588), 252-255.

Bramson, R., Sanders, C. W., Sadoski, M., West, C., Wiprud, R., English, M., Palm, M., & Xenakis, A. (2011). Comparing the effects of mental imagery rehearsal and physical practice on learning lumbar puncture by medical students. *Annals of Behavioral Science and Medical Education, 17*(2), 3-6.

Broadbent, D. P., Causer, J., Ford, P. R., & Williams, A. M. (2015). Contextual interference effect on perceptual-cognitive skills training. *Medicine and science in sports and exercise*, *47*(6), 1243-1250.

Brouziyne, M. & Molinaro, C. (2005). Mental imagery combined with physical practice of approach shots for golf beginners. *Perceptual and motor skills, 101*(1), 203-211.

Brydges, R., Carnahan, H., Backstein, D., & Dubrowski, A. (2007). Application of motor learning principles to complex surgical tasks: searching for the optimal practice schedule. *Journal of Motor Behavior*, *39*(1), 40-48.

Buchanan, J. J., & Wang, C. (2012). Overcoming the guidance effect in motor skill learning: feedback all the time can be beneficial. *Experimental brain research*, *219*(2), 305-320.

Burke, T. M., Scheer, F. A., Ronda, J. M., Czeisler, C. A., & Wright, K. P. (2015). Sleep inertia, sleep homeostatic and circadian influences on higher-order cognitive functions. *Journal of sleep research*, *24*(4), 364-371.

C

Carswell, C. M., Clarke, D., & Seales, W. B. (2005). Assessing mental workload during laparoscopic surgery. *Surgical innovation, 12*(1), 80-90.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin*, *132*(3), 354.

Chmarra, M. K., Bakker, N. H., Grimbergen, C. A., & Dankelman, J. (2006). TrEndo, a device for tracking minimally invasive surgical instruments in training setups. *Sensors and Actuators A: Physical, 126*(2), 328-334.

Chmarra, M. K., Dankelman, J., van den Dobbelsteen, J. J., & Jansen, F. W. (2008). Force feedback and basic laparoscopic skills. *Surgical endoscopy, 22*(10), 2140-2148.

Clark, R. E., Feldon, D., van Merriënboer, J. J., Yates, K., & Early, S. (2008). Cognitive task analysis. *Handbook of research on educational communications and technology*, *3*, 577-593.

Cook, D. A., Beckman, T. J., Thomas, K. G., & Thompson, W. G. (2008). Introducing resident doctors to complexity in ambulatory medicine. *Medical education*, *42*(8), 838-848.

Cotton, J. L. (1981). A review of research on Schachter's theory of emotion and the misattribution of arousal. *European Journal of Social Psychology*, *11*(4), 365-397.

Coyle, D. (2009). *The Talent Code: Greatness isn't born, it's grown. Here's How.* New York: Bantam Dell, a division of Ramdom House, Inc.

Crochet, P., Aggarwal, R., Dubb, S. S., Ziprin, P., Rajaretnam, N., Grantcharov, T., Ericsson, K. A., & Darzi, A. (2011). Deliberate practice on a virtual reality laparoscopic simulator enhances the quality of surgical technical skills. *Annals of surgery, 253(6)*, 1216-1222.

Cumming, J. L. & Ste-Marie, D. M. (2001). The cognitive and motivational effects of imagery training: A matter of perspective. *The Sport Psychologist, 15*, 276-288.

Cumming, J.L. & Hall, C. (2002): Deliberate imagery practice: the development of imagery skills in competitive athletes*. Journal of Sports Sciences, 20*(2), 137-145

D

Datta, V., Chang, A., Mackay, S., & Darzi, A. (2002). The relationship between motion analysis and surgical technical assessments. *The American journal of surgery, 184*(1), 70-73.

Dhand, R., & Sohal, H. (2006). Good sleep, bad sleep! The role of daytime naps in healthy adults. *Current opinion in pulmonary medicine*, *12*(6), 379-382.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*, 795-805.

Doyon, J., Korman, M., Morin, A., Dostie, V., Tahar, A. H., Benali, H., ... & Carrier, J. (2009). Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Experimental Brain Research*, *195*(1), 15-26.

Dreyfus, H. L., Dreyfus, S. E., & Athanasioum T. (1986). *Mind over Machine.* New York: Free Press.

Duffy, A. J., Hogle, N. J., McCarthy, H., Lew, J. I., Egan, A., Christos, P., & Fowler, D. L. (2005). Construct validity for the LAPSIM laparoscopic surgical simulator. *Surgical Endoscopy And Other Interventional Techniques, 19*(3), 401-405.

Duke, R. A., Allen, S. E., Cash, C. D., & Simmons, A. L. (2009). Effects of early and late rest breaks during training on overnight memory consolidation of a keyboard melody. *Annals of the New York Academy of Sciences*, *1169*(1), 169-172.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest, 14*(1), 4-58.

Dweck, C. (2006). *Mindset: The new psychology of success.* Random House Digital, Inc. Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, *32*(3), 158-172.

E

Ericsson, K. A. (2006). *The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance* (in 'The Cambridge handbook of expertise and expert performance' by Ericsson, K. A.)

Ericsson, K. A., Krampe, R.Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

F

Fadde, P.J. (2010). *Training Complex Psychomotor Performance Skills: A Part-Task Approach.* In: K. H. Silber & R. Foshay (eds.), Handbook of training and improving workplace performance. Volume 1: Instructional design and training delivery. For The International Society for Performance Improvement (New York: Pfeiffer of Wiley & Sons).

Fisherl, C. D. (1993). Boredom at work: A neglected concept. *Human Relations*, *46*(3), 395-417.

Friedlander, M. J., Andrews, L., Armstrong, E. G., Aschenbrenner, C., Kass, J. S., Ogden, P., Schwartzstein, R., & Viggiano, T. R. (2011). What can medical education learn from the neurobiology of learning?. *Academic Medicine*, *86*(4), 415-420.

G

Gallagher, A. G. & O'Sullivan, G. C. (2012). *Fundamentals of Surgical Simulation.* London: Springer.

Gallagher, A. G., Jordan-Black, J. A., & O'Sullivan, G. C. (2012). Prospective, randomized assessment of the acquisition, maintenance, and loss of laparoscopic skills. *Annals of surgery, 256*(2), 387-393.

Gallagher, A. G., McClure, N., McGuigan, J., Ritchie, K., & Sheehy, N. P. (2008). An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy, 30*(07), 617-620.

Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G., Smith, C. D. & Satava, R. M. (2005). Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Annals of surgery, 241*(2), 364.

Gauger, P. G., Hauge, L. S., Andreatta, P. B., Hamstra, S. J., Hillard, M. L., Arble, E. P., Kasten, S. J., Mullan, P. B., Cederna, P.S., & Minter, R. M. (2010). Laparoscopic simulation training with proficiency targets improves practice and performance of novice surgeons. *The American Journal of Surgery*, *199*(1), 72-80.

Goodell, K. H., Cao, C. G., & Schwaitzberg, S. D. (2006). Effects of cognitive distraction on performance of laparoscopic surgical tasks. *Journal of Laparoendoscopic & Advanced Surgical Techniques, 16*(2), 94-98.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills*. Acta Psychologica, 71*(1), 147-177.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, *37*(6), 504-528.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of motor behavior, 36*(2), 212-224.

H

Hallgató, E., Győri-Dani, D., Pekár, J., Janacsek, K., & Nemeth, D. (2012). The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex*, doi:10.1016/j.cortex.2012.01.002

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and motor skills*, *78*(3), 835-841.

Hall, C.R., Mack, D.E., Paivio, A., & Hausenblas, H.A. (1998). Imagery use by athletes: Development of the sport imagery questionnaire. *International Journal of Sport Psychology, 29*, 73-89.

Hamming, J. F. (2005). *Chirurg: teamspeler*. Universiteit Leiden, *Oratie tekst*. Derived from: https://www.lumc.nl/over-het-lumc/hoo/oraties-redes/2005/80806085042185/

Healy, A. F., & Bourne Jr, L. E. (Eds.). (2012). *Training cognition: Optimizing efficiency, durability, and generalizability*. Psychology Press.

Hiemstra, E. (2012). *Acquiring minimally invasive surgical skills*. Department of Minimally Invasive Surgery in Gynaecology, Faculty of Medicine/Leiden University Medical Center (LUMC), Leiden University.

Hiemstra, E., Kolkman, W., & Jansen, F. W. (2008). Skills training in minimally invasive surgery in Dutch obstetrics and gynecology residency curriculum. *Gynecological surgery, 5*(4), 321-325.

Hodges, B. D., & Kuper, A. (2012). Theory and practice in the design and conduct of graduate medical education. *Academic Medicine*, *87*(1), 25-33.

Horeman, T. (2014). *Force-based assessment of tissue handling skills*. TU Delft, Delft University of Technology.

Horeman, T., Blikkendaal, M. D., Feng, D., van Dijke, A., Jansen, F., Dankelman, J., & van den Dobbelsteen, J. J. (2014). Visual force feedback improves knot-tying security. *Journal of surgical education*, *71*(1), 133-141.

Horeman, T., Dankelman, J., Jansen, F. W., & van den Dobbelsteen, J. J. (2014). Assessment of laparoscopic skills based on force and motion parameters. *Biomedical Engineering, IEEE Transactions on*, *61*(3), 805-813.

Horeman, T., Rodrigues, S. P., van den Dobbelsteen, J. J., Jansen, F. W., & Dankelman, J. (2012). Visual force feedback in laparoscopic training. *Surgical endoscopy*, *26*(1), 242-248.

Horeman, T., van Delft, F., Blikkendaal, M. D., Dankelman, J., van den Dobbelsteen, J. J., & Jansen, F. W. (2014). Learning from visual force feedback in box trainers: tissue manipulation in laparoscopic surgery. *Surgical endoscopy*, *28*(6), 1961-1970.

Hsu, K. E., Man, F. Y., Gizicki, R. A., Feldman, L. S., & Fried, G. M. (2008). Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. *Surgical endoscopy, 22*(1), 196-201.

I

Immenroth, M., Bürger, T., Brenner, J., Nagelschmidt, M., Eberspächer, H., & Troidl, H. (2007). Mental training in surgical education: a randomized controlled trial. *Annals of surgery, 245*(3), 385.

Issa, N., Schuller, M., Santacaterina, S., Shapiro, M., Wang, E., Mayer, R. E., & DaRosa, D. A. (2011). Applying multimedia design principles enhances learning in medical education. *Medical education*, *45*(8), 818-826.

J

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*(11), 1419-1432.

Jones, L. L., & French, K. E. (2007). Effects of Contextual Interference on Acquisition and Retention of Three Volleyball Skills. *Perceptual and motor skills, 105*(3), 883-890.

Jordan, J. A., Gallagher, A. G., McGuigan, J., McGlade, K., & McClure, N. (2000). A comparison between randomly alternating imaging, normal laparoscopic imaging, and virtual reality training in laparoscopic psychomotor skill acquisition. *The American journal of surgery, 180(3)*, 208-211.

K

Kahol, K., Leyba, M. J., Deka, M., Deka, V., Mayes, S., Smith, M., Ferrara, J.J. & Panchanathan, S. (2008). Effect of fatigue on psychomotor and cognitive skills. *The American Journal of Surgery, 195*(2), 195-204.

Kassirer, J. P. (1976). The principles of clinical decision making: an introduction to decision analysis. *The Yale journal of biology and medicine*, *49*(2), 149.

Kolkman, W., Van de Put, M. A. J., Wolterbeek, R., Trimbos, J. B. M. Z., & Jansen, F. W. (2008). Laparoscopic skills simulator: construct validity and establishment of performance standards for residency training. *Gynecological Surgery, 5*(2), 109-114.

Kooloos, J. G., de Waal Malefijt, M. C., Ruiter, D. J., & Vorstenbosch, M. A. (2012). Loosely-guided, self-directed learning versus strictly-guided, station-based learning in gross anatomy laboratory sessions. *Anatomical sciences education*, *5*(6), 340-346.

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature neuroscience, 10*(9), 1206-1213.

Korman, M., Raz, N., Flash, T., & Karni, A. (2003). Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proceedings of the National Academy of Sciences*, *100*(21), 12492-12497.

Kuriyama, K., Stickgold, R., & Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory*, *11*(6), 705-713.

L

Lam, C. F., DeRue, D. S., Karam, E. P., & Hollenbeck, J. R. (2011). The impact of feedback frequency on learning and task performance: Challenging the “more is better” assumption. *Organizational Behavior and Human Decision Processes*, *116*(2), 217-228.

Lange, M. M., Buunen, M., van de Velde, C. J., & Lange, J. F. (2008). Level of arterial ligation in rectal cancer surgery: low tie preferred over high tie. A review. *Diseases of the colon & rectum*, *51*(7), 1139-1145.

Lee, T. D., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for exercise and Sport*, *59*(4), 277-287.

Leff, D. R., Aggarwal, R., Rana, M., Nakhjavani, B., Purkayastha, S., Khullar, V., & Darzi, A. W. (2008). Laparoscopic skills suffer on the first shift of sequential night shifts: program directors beware and residents prepare. *Annals of surgery*, *247*(3), 530-539.

Levinson, A. J. (2010). Where is evidence-based instructional design in medical education curriculum development?. *Medical education*, *44*(6), 536-537.

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G. R., Reznick, R., Bohnen, J., Orser, B., Doran, D., & Grober, E. (2004). Communication failures in the operating room: an observational classification of recurrent types and effects. *Quality and Safety in Health Care*, *13*(5), 330-334.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation On November 9th, 2011· In Business Models, Business Theories. *American psychologist, 57(9)*, 705-717.

Loehr, J. E., & Schwartz, T. (2003). *The Power of Full Engagement: Managing Energy, Not Time, Is the Key to High Performance and Personal Renewal.* Simon & Schuster, Inc.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological review, 95(4)*, 492-527.

London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management*, *32*(2-3), 353-372.

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research*, *24*(2), 199-205.

Luursema, J. M., Buzink, S. N., Verwey, W. B., & Jakimowicz, J. J. (2010). Visuo-spatial ability in colonoscopy simulator training. *Advances in health sciences education, 15*(5), 685-694.

Luursema, J. M., Verwey, W. B., & Burie, R. (2012). Visuospatial ability factors and performance variables in laparoscopic simulator training. *Learning and Individual Differences, 22*(5), 632-638.

Lynagh, M., Burton, R., & Sanson-Fisher, R. (2007). A systematic review of medical skills laboratory training: where to from here? *Medical Education, 41(9)*, 879-887.

M

Macedonia, M. (2002). Games, simulation, and the military education dilemma. *In Internet and the University: 2001 Forum* (pp. 157-167).

Magill, R. A., & Anderson, D. (2007). *Motor learning and control: Concepts and applications* (Vol. 11). New York: McGraw-Hill.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human movement science, 9*(3), 241-289.

Mané, A. M., Adams, J. A., & Donchin, E. (1989). Adaptive and part-whole training in the acquisition of a complex perceptual-motor skill. *Acta Psychologica, 71(1)*, 179-196.

Mann, K. V. (2011). Theoretical perspectives in medical education: Past experience and future possibilities. *Medical Education, 45*, 60-68.

Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery, 84(2)*, 273-278.

Masters, R.S., Lo, C.Y., Maxwell, J.P. & Patil, N.G. (2008). Implicit motor learning in surgery: implications for multi-tasking, *Surgery, 143*, 140-145.

Masters, R.S.W., MacMahon, K.M.A., & Pall, H.S. (2004). Implicit motor learning in Parkinson's disease. *Rehabilitation Psychology, 49*, 79-82.

Masters, R.S.W., Poolton, J.M., &Maxwell, J.P. et al., (2008). Implicit motor learning and complex decision making in time-constrained environments. *Journal of Motor Behavior, 40*, 71-79.

Maxwell, J.P., Masters, R.S.W., Kerr, E. et al. (2001). The implicit benefit of learning without errors. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 54*, 1049-1068.

McDaniel, M. (2012). *Put the SPRINT in Knowledge Training: Training with SPacing, Retrieval, and INTerleaving*. (in A. F. Healy & L. E. Bourne Jr. (Eds.), Training cognition: Optimizing Efficiency, Durability, and Generalizability (pp. 267-286).

McDermott, P. L., Carolan, T., & Wickens, C. D. (2012). Part Task Training Methods in Simulated and Realistic Tasks. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications, 56(1)*, 2502-2506)

McDougall, E. M., Corica, F. A., Boker, J. R., Sala, L. G., Stoliar, G., Borin, J. F., Chu, F. T. &Clayman, R. V. (2006). Construct validity testing of a laparoscopic surgical simulator. *Journal of the American College of Surgeons, 202*(5), 779-787.

McGaghie, W. C., Issenberg, S. B., Cohen, M. E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic medicine: journal of the Association of American Medical Colleges, 86(6)*, 706-711.

Miccoli, L., Versace, F., Koterle, S., & Cavallero, C. (2008). Comparing Sleep-Loss Sleepiness and Sleep Inertia: Lapses Make the Difference. *Chronobiology international*, *25*(5), 725-744.

Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*(11), 1618-1641.

Milner, C. E., Fogel, S. M., & Cote, K. A. (2006). Habitual napping moderates motor performance improvements following a short daytime nap. *Biological psychology*, *73*(2), 141-156.

Montgomery-Downs, H. E., Insana, S. P., & Bond, J. A. (2012). Movement toward a novel activity monitoring device. *Sleep and Breathing*, *16*(3), 913-917.

Moulton, C. A. E., Dubrowski, A., MacRae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Annals of surgery, 244*(3), 400.

Mullen, R., Hardy, L., & Oldham, A. (2007). Implicit and explicit control of motor actions: Revisiting some early evidence. *British Journal of Psychology*, 98, 141-156.

Mulrine, H. M., Signal, T. L., Berg, M. J. V. D., & Gander, P. H. (2012). Post-sleep inertia performance benefits of longer naps in simulated nightwork and extended operations. *Chronobiology international*, *29*(9), 1249-1257.

N

Nakamura, J., & Csikszentmihalyi, M. (2002). *The concept of flow.* Handbook of positive psychology, 89-105.

P

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4.

Patterson, J. T., Carter, M. J., & Hansen, S. (2013). Self-controlled KR schedules: Does repetition order matter?. *Human movement science*, *32*(4), 567-579.

Peck, A. C., & Detweiler, M. C. (2000). Training concurrent multistep procedural tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 42*(3), 379-389.

Peyre, S. E., Peyre, C. G., Hagen, J. A., Sullican, M. E., Lipham, J. C., DeMeester, S. R., Peters, J. H., & Demeester, T. R. (2009). Laparoscopic Nissen fundoplication assessment: task analysis as a model for the development of a procedural checklist. *Surgical endoscopy*, *23*(6), 1227-1232.

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30(1)*, 96-116.

Pluyter, J. R., Buzink, S. N., Rutkowski, A. F., & Jakimowicz, J. J. (2010). Do absorption and realistic distraction influence performance of component task surgical procedure?. *Surgical endoscopy, 24*(4), 902-907.

Pluyter, J. R., Rutkowski, A., Jakimowicz, J. J., & Saunders, C. S. (2012). Measuring Users' Mental Strain When Performing Technology Based Surgical Tasks on a Surgical Simulator Using Thermal Imaging Technology. *45th Hawaii International Conference on System Science (HICSS)*, (pp. 2920-2926).

Poolton, J.M., Masters, R.S.W., & Maxwell, J.P. (2007). Passing thoughts on the evolutionary stability of implicit motor behavior: Performance retention under physiological fatigue. *Consciousness & Cognition*, 16, 456-468.

Porter, J. M., & Magill, R. A. (2010). Systematically increasing contextual interference is beneficial for learning sport skills. *Journal of sports sciences*, *28*(12), 1277-1285.

Prasad, S. M., Prasad, S. M., Maniar, H. S., Chu, C., Schuessler, R. B., & Damiano, R. J. (2004). Surgical robotics: impact of motion scaling on task performance. *Journal of the American College of Surgeons*, *199*(6), 863-868.

Proteau, L., Blandin, Y., Alain, C., & Dorion, A. (1994). The effects of the amount and variability of practice on the learning of a multi-segmented motor task. *Acta psychologica, 85*(1), 61-74.

Purves, D., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L. & Woldorff, M. G. (2012). *Principles of Cognitive Neuroscience, Second Edition*. Sinauer Associates, Inc.

R

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 5, 855-863.

Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills - changes in the wind. *New England Journal of Medicine, 355*(25), 2664-2669.

Ribeiro, D. C., Sole, G., Abbott, J. H., & Milosavljevic, S. (2011). Extrinsic feedback and management of low back pain: A critical review of the literature. *Manual therapy*, *16*(3), 231-239.

Rieth, C. A., Cai, D. J., McDevitt, E. A., & Mednick, S. C. (2010). The role of sleep and practice in implicit and explicit motor learning. *Behavioural brain research*, *214*(2), 470-474.

Risucci, D., Geiss, A., Gellman, L., Pinard, B., & Rosser, J. (2001). Surgeon-specific factors in the acquisition of laparoscopic surgical skills. *The American Journal of Surgery, 181*(4), 289-293.

Ritter, E. M., McClusky, D. A., Gallagher, A. G., & Smith, C. D. (2005). Real-time objective assessment of knot quality with a portable tensiometer is superior to execution time for assessment of laparoscopic knot-tying performance. *Surgical innovation, 12(3)*, 233-237.

Rodrigues, S. P., Horeman, T., Dankelman, J., van den Dobbelsteen, J. J., & Jansen, F. W. (2012). Suturing intraabdominal organs: when do we cause tissue damage?. *Surgical endoscopy*, *26*(4), 1005-1009.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, *39*(5), 406-412.

Rook, J. W., & Zijlstra, F. R. (2006). The contribution of various types of activities to recovery. *European journal of work and organizational psychology, 15*(2), 218-240.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual review of psychology*, *52*(1), 453-470.

S

SAGES: Society of American Gastrointestinal and Endoscopic Surgeons (2013). Fundamentals of Laparoscopic Surgery (FLS). http://www.flsprogram.org/wp-content/uploads/2013/05/FLS-Information-Bulletin-May-2013.doc

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The international journal of aviation psychology, 8*(3), 197-208.

Salden, R. J., Paas, F., & van Merrienboer, J. J. (2006). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Computers in Human Behavior, 22*(3), 321-333.

Salkini, M. W., & Hamilton, A. J. (2010). The effect of age on acquiring laparoscopic skills. *Journal of Endourology, 24*(3), 377-379.

Sanders, C. W., Sadoski, M., Bramson, R., Wiprud, R., & Van Walsum, K. (2004). Comparing the effects of physical practice and mental imagery rehearsal on learning basic surgical skills by medical students. *American journal of obstetrics and gynecology, 191*(5), 1811-1814.

Sanders, C. W., Sadoski, M., van Walsum, K., Bramson, R., Wiprud, R., & Fossum, T. W. (2008). Learning basic surgical skills with mental imagery: using the simulation centre in the mind. *Medical education, 42*(6), 607-612.

Sanders, C. W., Sadoski, M., Wasserman, R. M., Wiprud, R., English, M., & Bramson, R. (2007). Comparing the effects of physical practice and mental imagery rehearsal on learning basic venipuncture by medical students*. Imagination, Cognition and Personality, 27*(2), 117-127.

Schaverien, M. V. (2010). Development of expertise in surgical training. *Journal of surgical education*, *67*(1), 37-43.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological science, 3*(4), 207-217.

Schneider, V. I., Healy, A. F., & Bourne Jr, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, *46*(2), 419-440.

Schols, R. M., Dunias, P., Wieringa, F. P., & Stassen, L. P. (2013). Multispectral characterization of tissues encountered during laparoscopic colorectal surgery. *Medical engineering & physics*, *35*(7), 1044-1050.

Schreuder, H. W., Oei, G., Maas, M., Borleffs, J. C., & Schijven, M. P. (2011). Implementation of simulation in surgical practice: minimally invasive surgery has taken the lead: the Dutch experience. *Medical Teacher, 33*(2), 105-115.

Seijts, G. H., & Latham, G. P. (2006). Learning goals or performance goals: is it the journey or the destination*? Ivey Business Journal*, May/June 2006: 1-6.

Seymour, N. E. (2008). VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World journal of surgery, 32*(2), 182-188.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery, 236*(4), 458.

Shapiro, M. J., Morey, J. C., Small, S. D., Langford, V., Kaylor, C. J., Jagminas, L., Suner, S., Salisbury, M. L., Simon, R. & Jay, G. D. (2004). Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum?. *Quality and Safety in Health Care, 13*(6), 417-421.

Shea, J. B., & Graf, R. C. (1994). A model for contextual interference effects in motor learning. In C. R. Reynolds (Ed.), *Cognitive assessment: a multidisciplinary perspective, 73-87.* New York: Plenum Press.

Shea, C. H., Lai, Q., Black, C., & Park, J. H. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human movement science, 19*(5), 737-760.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179−187.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological review, 84(2)*, 127-190.

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review*, *20*(1), 21-53.

Sinitsky, D. M., Fernando, B., & Berlingieri, P. (2012). Establishing a curriculum for the acquisition of laparoscopic psychomotor skills in the virtual reality environment. *The American Journal of Surgery*, *204*(3), 367-376.

Smink, D. S., Peyre, S. E., Soybel, D. I., Tavakkolizadeh, A., Vernon, A. H., & Anastakis, D. J. (2012). Utilization of a cognitive task analysis for laparoscopic appendectomy to identify differentiated intraoperative teaching objectives. *The American Journal of Surgery*, *203*(4), 540-545.

Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology Learning Memory and Cognition, 30*(3), 601-604.

Spruit E.N., Band G.P.H. & Hamming J.F. (2013), Feedback bij het aanleren van laparoscopie, *Nederlands Tijdschrift voor Geneeskunde* 157(23): A6354

Spruit, E. N., Band, G. P., & Hamming, J. F. (2015). Increasing efficiency of surgical training: effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surgical endoscopy*, *29*(8), 2235-2243.

Spruit, E. N., Band, G. P., Hamming, J. F., & Ridderinkhof, K. R. (2014). Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. *Psychological research*, *78*(6), 878-891.

Spruit, E. N., Kleijweg, L. M., Band, G. P., & Hamming, J. F. (2016). Varied Practice in Laparoscopy Training: Beneficial Learning Stimulation or Cognitive Overload?. *Frontiers in Psychology*, *7*, 685.

Stefanidis, D., & Heniford, B. T. (2009). The formula for a successful laparoscopic skills curriculum. *Archives of Surgery*, *144*(1), 77-82.

Stefanidis, D., Korndorffer, J. R., Markley, S., Sierra, R., Heniford, B. T., & Scott, D. J. (2007). Closing the gap in operative performance between novices and experts: does harder mean better for laparoscopic simulator training?. *Journal of the American College of Surgeons*, *205*(2), 307-313.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, *437*(7063), 1272-1278.

Strandbygaard, J., Bjerrum, F., Maagaard, M., Winkel, P., Larsen, C. R., Ringsted, C., Gluud, C., Grantcharov, T., Ottesen, B. & Sorensen, J. L. (2013). Instructor Feedback Versus No Instructor Feedback on Performance in a Laparoscopic Virtual Reality Simulator. *Annals of Surgery 257*(5), 839–844.

Sturm, L. P., Windsor, J. A., Cosman, P. H., Cregan, P., Hewett, P. J., & Maddern, G. J. (2008). A systematic review of skills transfer after surgical simulation training. *Annals of surgery, 248*(2), 166-179.

Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: a review for its modernization. *Anatomical sciences education*, *3*(2), 83-93.

Sullivan, M. E., Ortega, A., Wasserberg, N., Kaufman, H., Nyquist, J., & Clark, R. (2008). Assessing the teaching of procedural skills: can cognitive task analysis add to our traditional teaching methods?. *The American Journal of Surgery*, *195*(1), 20-23.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review, 10(3)*, 251-296.

T

Taber, M. J. (2014). Simulation fidelity and contextual interference in helicopter underwater egress training: An analysis of training and retention of egress skills. *Safety science*, *62*, 271-278.

Teague, R. C., Gittelman, S. S., & Park, O. C. (1994). *A review of literature on part-task and whole-task training and context dependency.* Alexandria, VA: Army Research Institute for the Behavioral and Social Social Sciences.

Thijssen, A. S., & Schijven, M. P. (2010). Contemporary virtual reality laparoscopy simulators: quicksand or solid grounds for assessing surgical trainees? *The American journal of surgery, 199(4)*, 529-541.

Tholey, G., Desai, J. P., & Castellanos, A. E. (2005). Force feedback plays a significant role in minimally invasive surgery: results and analysis. *Annals of surgery, 241*(1), 102-109.

U

Undre, S., Sevdalis, N., Healey, A. N., Darzi, A., & Vincent, C. A. (2007). Observational teamwork assessment for surgery (OTAS): refinement and application in urological surgery. *World journal of surgery*, *31*(7), 1373-1381.

V

Van den Dobbelsteen, J. J., Schooleman, A., & Dankelman, J. (2007). Friction dynamics of trocars. *Surgical endoscopy*, *21*(8), 1338-1343.

Van der Linden, D. (2011). *The Urge to Stop: The Cognitive and Biological Nature of Acute Mental Fatigue.* (in P. L. Ackerman (Ed.). Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications (pp. 149-164)).

Van der Wal, G. (2007). *Risico's minimaal invasieve chirurgie onderschat.* Inspectie voor de Gezondheidszorg, Den Haag, NL.

Van Det, M. J., Meijerink, W. J. H. J., Hoff, C., Totte, E. R., & Pierie, J. P. E. N. (2009). Optimal ergonomics for laparoscopic surgery in minimally invasive surgery suites: a review and guidelines. *Surgical endoscopy*, *23*(6), 1279-1285.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*(1), 16-26.

Van Gog, T., Ericsson, K. A., Rikers, R. M., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development, 53(3)*, 73-81.

van Hove, P. D., Verdaasdonk, E. G., van der Harst, E., Jansen, F. W., Dankelman, J., & Stassen, L. P. (2012). Effect of basic laparoscopic skills courses on essential knowledge of equipment. *Surgical innovation*, *19*(4), 460-466.

Van Merriënboer, J. J., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks:Balancing intrinsic and germane load to enhance transfer of learning. *Applied cognitive psychology*, *20*(3), 343-352.

Van Merriënboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical education*, *44*(1), 85-93.

Van Sickle, K. R., Baghai, M., Huang, I. P., Goldenberg, A., Smith, C. D., & Ritter, E. M. (2008). Construct validity of an objective assessment method for laparoscopic intracorporeal suturing and knot tying. *American journal of surgery, 196*(1), 74-80.

Våpenstad, C., Hofstad, E. F., Langø, T., Mårvik, R., & Chmarra, M. K. (2013). Perceiving haptic feedback in virtual reality simulators. *Surgical endoscopy*, *27*(7), 2391-2397.

W

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron*, *35*(1), 205-211.

Wang, Z., Zhou, R., & Shah, P. (2014). Spaced cognitive training promotes training transfer. *Frontiers in human neuroscience*, *8*, 217.

Wherry, D. C., Rob, C. G., Marohn, M. R., & Rich, N. M. (1994). An external audit of laparoscopic cholecystectomy performed in medical treatment facilities of the Department of Defense. *Annals of surgery*, *220*(5), 626.

Whiting, H. T. A., & den Brinker, B. P. (1981). *Image of the act.* (in 'Learning difficulties' by Das, J. P., Mulcahy, R., & Wall, A. E. (Eds.).

Wickens, C D, Hutchins, S, Carolan, T, & Cumming, J. (2013). Effectiveness of part-task training and increasing-difficulty training strategies: A meta-analysis approach. *Human Factors, 55*(2), 461-470.

Wickens, C., Hutchins, S., Carolan, T., & Cumming, J. (2011). Investigating the Impact of Training on Transfer A Meta-Analytic Approach. *SAGE Publications: In Proceedings of the Human Factors and Ergonomics Society Annual Meeting 55*(1), 2138-2142.

Wightman, D.C. & Lintern, G. (1985). Part-task training for tracking and manual control. *Human Factors, 27*, 267-283.

Wiklund, N. P. (2004). Technology Insight: surgical robots—expensive toys or the future of urologic surgery?. *Nature Clinical Practice Urology*, *1*(2), 97-102.

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science, 18*(2), 133-134.

Wong, S. W., Smith, R., & Crowe, P. (2010). Optimizing the operating theatre environment. *ANZ journal of surgery*, *80*(12), 917-924.

Wottawa, C. R., Cohen, J. R., Fan, R. E., Bisley, J. W., Culjat, M. O., Grundfest, W. S., & Dutson, E. P. (2013). The role of tactile feedback in grip force during laparoscopic training tasks. *Surgical endoscopy*, *27*(4), 1111-1118.

Wulf, G. (2007). *Attention and motor skill learning*. Human Kinetics.

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: a review of influential factors. *Medical education, 44*(1), 75-84.

Z

Zajonc, R. B. (1965). *Social facilitation*. Research Center for Group Dynamics, Institute for Social Research, University of Michigan

Zendejas, B., Brydges, R., Hamstra, S. J., & Cook, D. A. (2013). State of the Evidence on Simulation-Based Training for Laparoscopic Surgery: A Systematic Review. *Annals of surgery, 257*(4), 586-593.

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: A design approach for modern tools*. TU Delft, Delft University of Technology.

CHAPTER 12

APPENDIX A

SUB-TASKS, OPERATIVE STEPS AND DECISION POINTS DURING LAPAROSCOPIC SIGMOID COLON RESECTION

Appendix A. Sub-tasks, operative steps and Decision points during Laparoscopic Sigmoid Colon Resection

Sub-task	Operative steps & *Decision Points* (N): Novice, (I): Intermediate, (A): Advanced
Patient positioning & Trocar placement	• (N) Time-out procedure
	• (N) Place patient in Trendelenburg with a 15-25 degree vertical tilt and a 5-10 degree tilt to the right. Legs are spread with feet at the level of the hips.
	• (N) Surgeon and assisting surgeon are positioned to the right of the patient
	• (I) Place first trocar at the umbilicus
	• (N) Place second trocar at the right flank
	• (N) Place third trocar just above the pubic bone
	• (I) *Given the patients size, are more trocars needed?* *Yes* -> Place fourth/fifth/sixth trocar *No* -> Proceed to next sub-task
Investigation of the abdominal cavity	• (I) Check for spread of infection to other organs
	• (I) Move the omentum and the small intestines anteriorly to facilitate exposure of the sigmoid colon
	• (I) Investigate the colic tumor locally
	• (A) *Is it an ingrown tumor?* *No* -> proceed to next sub-task *Yes -> grown into which structures?* *Left ureter* -> Convert to open (laparotomy) *Abdominal wall / bladder / appendix / coecum / small intestine* *-> Is adequate visualization possible?* *Yes* -> Proceed to next task *No* -> Convert to open
	• (I) *Is there adequate visualization of the sigmoid mesocolon and the trunk of the sigmoid arteries?* *Yes* -> Adopt Medial-to-Lateral approach *No* -> Mobilize lateral attachments first
Creating a submesentaric window & transecting the sigmoid veins	• (I) Suspend the uterus to facilitate exposure (optional in the case of post-menopausal female patients)
	• (I) Identify the sigmoid mesocolon
	• (I) *Where is the tumor located exactly?* *Sigmoid -> Dissect sigmoid arteries* *Colosigmoid -> Dissect sigmoid arteries and colica sinistra* *Rectosigmoid -> Dissect sigmoid arteries, rectus superior and the inferior mesentaric artery just above the branch of the Sigmoid Arteries*
	• (I) Establish rough margins for the to-be-extracted colon specimen at five centimeters distal and proximal to the tumor (while also taking into account blood supply) in order to guide the dissection of the sigmoid mesocolon and the dissection of the artery supply
	• (I) Apply traction ventrally to facilitate identification of the trunk of sigmoid arteries
	• (I) Blunt dissection of the avascular layers of the sigmoid mesocolon to create a submesentaric window

Sub-task	Operative steps & *Decision Points* (N): Novice, (I): Intermediate, (A): Advanced
Creating a submesentaric window & transecting the sigmoid veins	• (A) Expose and identify the trunk of the sigmoid arteries • (A) Expose and identify the left ureter • (A) Expose and identify the hypo gastric nerve • (A) Transect the sigmoid veins at a level that includes a minimum of twelve lymph nodes attached to the to-be-extracted colon specimen using a stapler, clips or Ligasure • (A) Transect Rectus Superior, Inferior Mesentaric Artery or Colica Sinistra if necessary • (I) Further dissect the peritoneum of the sigmoid mesocolon and lift it ventrally while preserving the lymph nodes
Posterior and lateral mobilization of the Sigmoid Colon	• (I) Dissect the posterior attachments of the sigmoid colon to the retroperitoneum • (I) Dissect the lateral attachments of the sigmoid colon to the abdominal wall • (A) *Is there a tensionless fit of the upper and lower end of the colon?* *Yes* -> Proceed to Division of the sigmoid colon *No* -> Proceed to Mobilization of the left colon towards the splenic flexure
Mobilization of the descending colon towards the splenic flexure	• (A) Mobilize the descending colon (and splenic flexure) laterally to gain length on the proximal colon • (A) *Is there a tensionless fit of the upper and lower end of the colon?* *Yes* -> Proceed to Division of the sigmoid colon *No* -> Mobilize the descending colon (and splenic flexure) medially to gain length on the proximal colon
Division of the sigmoid colon	• (A) Divide the distal sigmoid colon with a stapler at a minimum of five centimeters below the tumor • (I) Create a 6-centimer incision above the pubic bone • (I) Extract the sigmoid colon using a wound protector • (I) Divide the proximal mesocolon with a stapler at the level of the division of the sigmoid arteries • (I) Divide the proximal sigmoid colon with a stapler at a minimum of five centimeters above the tumor
Creating the anastomosis	• (A) *What kind of anastomosis will be created?* *End-to-end, side-to-side or side-to-end* • (I) Place and suture the first anvil in the proximal end of the colon • (I) Put the proximal colon back into the abdominal cavity • (I) Re-establish the pneumoperitoneum • (I) Place the circular stapler in the distal end of the colon/rectum via the anus • (I) Connect the anvil to the stapler • (I) Check that the proximal colon is not twisted • (N) Fire the circular stapler • (I) Optional: Create a stoma • (A) Test the anastomosis for leakage • (A) *Given the current dissection, is there high risk for inguinal hernia? Yes ->* Close the sigmoid mesocolon • (I) Final inspection and cleaning of the operative field • (I) Close the trocar wounds

CHAPTER 13

APPENDIX B

SURGICAL SKILLS PROFILE

Appendix B. Surgical Skills Profile

Form of Proficiency	Surgical Skill	Relevance
Surgical knowledge	Anatomy (including anomalies)	4
	Procedural steps	5
	Pathology	4
Perceptual skills	Identification of tissue types (anatomy)	5
	Depth perception	5
	Haptics	5
	Mental Imagery	5
Motor skills	Instrument handling	5
	Blunt dissection	5
	Mobilizing tissue for exposure/dissection	5
	Sharp dissection	5
	Using cautery/ultrasonic shears/stapler/etc.	5
	Intra-corporeal suturing	3
Clinical decision making	Switching operative techniques	4
	Changing pace of operation	4
	Situation awareness	4
	Meta cognition of skill acquisition	4
	Performance monitoring	3
	Response inhibition	3
Mental endurance	Emotional stability	4
	Mental focus (concentration)	4
Social skills	Team communication	3
	Instructing assisting staff	3
Technology skills	Understanding the mechanics of hard-ware	3
	Trouble-shooting equipment	3

CHAPTER 14

ACKNOWLEDGEMENTS

Acknowledgements

First, I would like to thank my promoters and supervisors, Guido Band, Jaap Hamming and Bernhard Hommel for providing me with the opportunity to work on this research project. I appreciated the autonomy, encouragement and constructive feedback they provided to me. At first, the autonomy was a bit overwhelming, but with time, all of that became a lot easier as we managed to establish a more specified focus in which the research project would be going. In the private sphere, I first and foremost want to thank my mother and father, Truda and Ed Sr., who have always supported me throughout my studies and research years. I want to thank my lovely wife, Ana, with whom I could always exchange experiences about the good and the bad of both of our PhD years and who would be waiting for me with a delicious warm meal after work, even late at midnight after returning from the lab. I also want to thank her for being the occasional lab assistant just to help out and see what I'm working on, providing feedback on each other's papers and presentations and just being such a joy to all people around her. Then I would like to thank my friends Koen, Arian, Tim, Rick, Jeroen and the extended skateboarding community in Dordrecht for the skating and chilling sessions as a welcome escape from work. I also want to thank my sisters, Ellen and Eva, their husbands, Martin and Arie and their children (Anouk, Hugo, Tijs and Fleur), as well as my extended family for simply having a good time together. Special thanks goes to Gertjan Hultzer and Rene Rodenburg, skillslab coordinators at the LUMC, for always bringing a good vibe and making sure everything in the lab is in order. I'd like to thank all the colleagues over at the cognitive psychology unit (Sander Nieuwenhuis, Pascal Haazebroek, William Verschuur, Mikael Kowal, Kerwin Olfers, Rudy van den Brink, Roy de Kleijn, Henk van Steenbergen, Roderik Gerritsen, Ke Ma, Wido la Heij, Roberta Sellaro, Klodiana-Daphne Tona, George Kachergis, Bruno Bocanegra, Peter Murphy and Christopher Warren) for providing feedback on my studies, the mandatory table Fußball, with a special shout out to the passionate calligraphic and videographic artwork by Stephen Brown, as well as Lorenza Colzato for providing the opportunity to write for the LIBC blog. I'd like to thank Dr. Hueting for providing me the opportunity to have an in-depth look at several laparoscopic interventions. I want to thank Tim Horeman for the pleasant cooperation during the collaboration study described in this dissertation, as well as providing us with the simulation models to perform it. I'd like to thank all of the students that participated in my studies, as well as the ones that I've supervised, with special mention to those I've had the opportunity to work with on the chapters included in this dissertation, Tom, Daan and Luca, supervising students with solid initiative truly makes teaching worthwhile. Finally, I want to thank Sidney Latuperissa for providing the wonderful artwork on the cover.

CHAPTER 15

CURRICULUM VITAE AND PUBLICATION LIST

Curriculum Vitae & Publication list

Personalia

Name:	Spruit
First names:	Edward Nicolaas
Date of birth:	25-03-1987
E-mailadress:	edwardspruit@hotmail.com

Education

2012 – 2016	PhD research project in the fields of laparoscopic surgery and applied cognitive psychology
2013	EPOS Graduate Education Network, Egmond aan Zee, Positive Journal Interaction Course
2009 – 2011	Fuentes, Dordrecht, Spanish for beginners 1 & 2; Conversation course A1, Spanish for semi-intermediates 1 & 2; Conversation course A2; Spanish for intermediates 1 & 2
2009 – 2011	Erasmus Universiteit, Rotterdam, Master Organizational Psychology (degree received in november 2011)
2009	Erasmus Universiteit, Rotterdam, Michael Jensen Leadership Seminar
2006 – 2010	Erasmus Universiteit, Rotterdam, Bachelor Organizational Psychology (degree received in may 2010)
2004 – 2006	Walburg College, Zwijndrecht, VWO Nature & Technology
1999 – 2004	Walburg College, Zwijndrecht, HAVO Nature & Technology

Work experience

2017	Rostock University Medical Center, Clinic for Psychiatry and Psychotherapy, Scientific employee / Psychologist. 50% clinical work on a qualified detoxification ward (alcohol and drug addiction) 50% research activities on the topics: resilience to psychiatric disorders affecting the autonomic nervous system (PTSD, anxiety disorders) Work activities: Conducting research activities, write and publish scientific papers, psycho-diagnostic testing, report writing, group and individual counseling, team meetings.

2012 - 2016 Leiden University, Leiden University Medical Center,
Department of Surgery & Applied Cognitive Psychology, PhD candidate.

Project title:
'Increasing the efficiency of laparoscopy training'.
The aim of this project was to improve training of medical staff (specifically surgical trainees in laparoscopy) by testing and implementing interventions in order to optimize consolidation of perceptual and motor skills.

Work activities:
Design, lead, execute and coordinate research plans. Write research proposals, perform empirical studies, instruct and supervise research assistants, analyze data, write and publish empirical papers, supervise students in internships and bachelor/masterthesis, give training in laparoscopy, perform semi-structured interviews with surgeons, cognitive task analysis, teaching in applied cognitive psychology, statistics and medical training, perform verbal and poster presentations on conferences, inter collogue feedback, write dissertation.
(http://www.narcis.nl/research/RecordID/OND1349342)

2010 April-June **Internship at EC Rijksadvies** (ministry of internal affairs), an organizational consultancy company for the government. The advisors of Rijksadvies support clients within departments of the government on a broad scale of organizational decisions. During the internship I cooperated on two projects with Annemiek Mul, organizational consultant at Rijksadvies. The first project was a pilot for Dienst Justitiële Instellingen (Justice). The other project was an evaluation task about the VAMEX organization for Verkeer en Waterstaat (ministry of transport).

Work activities:
Record, report and organize information, think and reflect with project groups, small secretarial tasks, develop documents to lead semi-structured interviews, participate in interviews and mini-conferences, write evaluation report (end product) and participation in organizational meetings.

1998 – 2011 Multiple side jobs (paper route, sales & construction employee stores, restaurant, production work).

Bachelorthesis

Literature review on the relationship between explanatory style and goal setting theory.

Masterthesis

Experiment into the relationship between mental fatigue, pupil responses and performance.
(http://identityisdynamic.com/2011/12/06/mental-fatigue-pupil-responses-and-performance/)

Publications

Spruit E.N., Band G.P.H. & Hamming J.F. (2013), Feedback bij het aanleren van laparoscopie, Nederlands Tijdschrift voor Geneeskunde 157(23): A6354.

Spruit, E.N., Band, G. P., Hamming, J. F., & Ridderinkhof, K. R. (2014). Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. *Psychological research*, *78*(6), 878-891.

Spruit, E.N., Band, G. P., & Hamming, J. F. (2015). Increasing efficiency of surgical training: effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surgical endoscopy*, *29*(8), 2235-2243.

Spruit E.N., Kleijweg L., Band G.P.H. & Hamming J.F. (2016). Varied Practice in Laparoscopy Training: Beneficial Learning Stimulation or Cognitive Overload?. Frontiers in Psychology, 7, 685.

Smit, D., Spruit, E., Dankelman, J., Tuijthof, G., Hamming, J., & Horeman, T. (2017). Improving training of laparoscopic tissue manipulation skills using various visual force feedback types. *Surgical endoscopy*, *31*(1), 299-308.

Spruit, E.N., Band, G. P., van der Heijden, K. B., & Hamming, J. F. (2017). The Effects of Spacing, Naps, and Fatigue on the Acquisition and Retention of Laparoscopic Skills. *Journal of surgical education*, *74*(3), 530-538.

Spruit, E.N., Colla, M., Gertz, K., & Kronenberg, G. (2017). Psychological Mechanisms Undergirding Resilience after Trauma. (resubmitted after first request for revision)

Hobby's/Interests

Skateboarding | Strength Training and Mobility | Yoga | Writing
(online book available at my website: www.identityisdynamic.com)

<u>Personal development (books, audio/video, courses, seminars):</u>

Success, performance and effectiveness | Social dynamics, social psychology | Physical, emotional and mental health (nutrition, fitness, meditation, etc.)

CHAPTER 16

ABOUT THE AUTHOR

About the Author

Edward N. Spruit was born March 25th 1987 in Dordrecht, the Netherlands. After completing "Hoger Algemeen Voortgezet Onderwijs" and "Voorbereidend Wetenschappelijk Onderwijs" at Walburg College in Zwijndrecht (2006) he commenced his Bachelor studies in Psychology at the Erasmus University Rotterdam. He obtained his Master of Science degree in Organizational Psychology in 2011. In 2012, he started his PhD research at Leiden University in a collaboration project between the department of Cognitive Psychology and Surgery, supervised by Prof. J. F. Hamming and Dr. G. P. H. Band, as described in this dissertation. Currently (2017), he works as a scientific employee in Rostock University Medical Center, focusing on the topic of psychological and physiological resilience to psychiatric disorders.